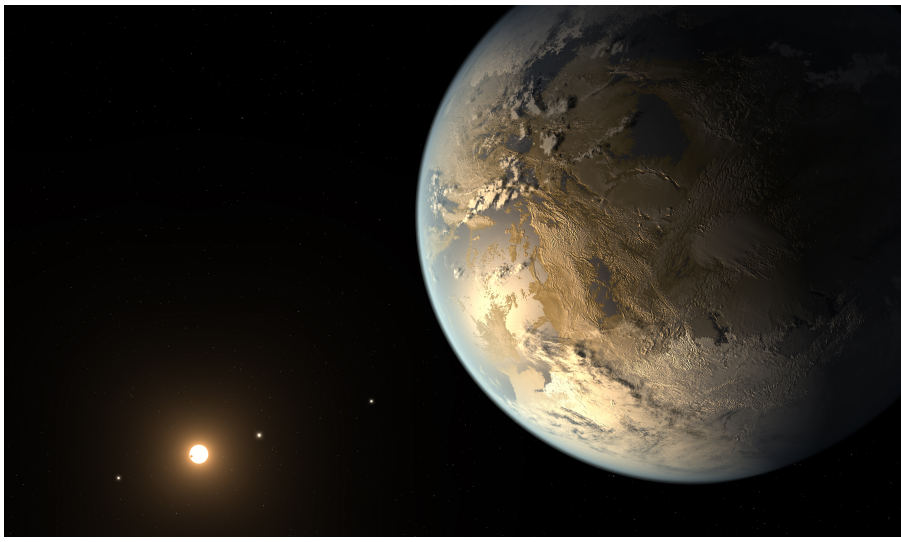


Exoplanet detection in metal-poor stars

João Pedro de Sousa Faria
Program Doutoral em Astronomia
Departamento de Física e Astronomia
2018

Orientador
Nuno M. C. Santos, Professor Auxiliar, Faculdade de Ciências

Coorientador
Pedro R. L. Figueira, Investigador Auxiliar, Instituto de Astrofísica e Ciências do Espaço

COLOPHON

This document was typeset in LaTeX, using modified versions of the `tufte-book` and `classicthesis` classes. Hermann Zapf's Palladio acts as the typeface throughout. The bibliography was processed with `biblatex`. Thank you to Aaron Turon for sharing his LaTeX code.

Title page illustration: NASA Ames / JPL-Caltech / T. Pyle

Acknowledgements

I'm supposed to thank all the people that made my PhD possible. And we'll go into that in just a bit, but first I want to acknowledge the faceless but enduring financial support from Fundação para a Ciência e a Tecnologia (FCT), in the form of a PhD studentship with reference SFRH/BD/93848/2013. I am just as indebted to every other grant I received during my PhD, but that list is lengthy.

Nuno and Pedro were the best supervisors I could have asked for; if I'm any good as a scientist after these four years, it's because of them. They provided me with so much: one incredible dataset to play with, guidance when I was lost, freedom when I (thought I) knew the way, and calm when the world infuriated me. I will be forever in debt.

Sometimes I wonder how many people stopped by my (always messy) desk in room 1.03. Pretty sure we're in the several dozens by now. And a few of them are my friends! Pedro (with the supervisor hat off, in good and not so good moods, but always with friendly advice and unconditional support), Ângela, Mahmoud, Vardan, Annelies, Elisa, Daniel, Jorge, more friends I forgot because I have a terrible memory, thank you for the fun moments we shared. It's also incredible the number of people I was fortunate enough to meet, either when they travelled here or I travelled there. Raphie, thank you for inspiring me, and for every other thing we got to do together. Thank you Rodrigo, Mario, Aldo, Alessandro, Xavier, Brendon, the cool people at La Silla. Over here at the institute, thank you Paulo for your help in hacking the Geneva database (a very harmless hack with an actual password, but still), thank you Nelma for all the help when I travelled and how easy you made everything look, and Argentina, Elsa, and Manuel. Thank you Mário João, for being an exponent of scientific integrity and an inspiration.

Making mine the words of others, since the feeling is just the same; in the vastness of space and immensity of time, it is my joy to spend a planet and an epoch with Solène.

Finally, the most important and hardest to put into words. I would not be writing this today were it not for my mother and all the sacrifices she endured over the years. I never know exactly how to say it, and end up not doing it often enough, but you are the reason why this is all possible, and I love you very much. This thesis is for you.

Abstract

Only two decades after the first detections of extra-solar planets, we are now taking sure steps towards finding planets similar to the Earth. In the mean time, the number of super-Earths and Neptune-like planets is growing, and the properties and frequency of such systems are being revealed. Planet formation models suggest that these low-mass planets, unlike giant planets, should be frequent around stars with low metallicities. In this Thesis, I test these model predictions by studying a sample of moderately metal-poor stars in the search for planetary companions.

I present and develop upon a method for the analysis of radial-velocity data, with which we are able to simultaneously infer the orbital parameters of the planets and the actual number of planets confidently detected in a given dataset. I also describe a rigorous criterion for planet detection, based on the amount of evidence supporting the presence of Keplerian signals in the data. The method is further extended by considering Gaussian processes as flexible models for the correlated noise that arises from stellar-induced radial-velocity variations. These variations, when not treated, can easily mimic or hide true radial-velocity signals of Earth-like planets. By modelling the correlated noise, I also manage to obtain information about the star and the magnetic activity at its surface.

The application of this model to the CoRoT-7 planetary system allowed for the confident detection of two low-mass planets, and also provided constraints on the stellar rotation period and the timescale for evolution of active regions, all obtained solely from radial-velocity observations. With the same method, I do not find enough evidence for the reported planet detections around HD 175607 and HD 41248, which are part of our metal-poor sample.

After analysing the full sample of 109 metal-poor stars, I find evidence for a significantly lower frequency of low-mass planets, when compared with results for stars with solar metallicity. These results challenge the predictions of the standard core-accretion theory, and strengthen the idea that planets with lower masses which form in metal-poor disks may not migrate far before the disk dissipates, ending up at larger orbital periods and being still undetectable.

Resumo

Apenas duas décadas após as primeiras detecções de planetas extra-solares, caminhamos a passos largos para encontrar planetas como a Terra. Entretanto, o número de super-Terras e Neptunos continua a crescer, e as propriedades e frequência destes sistemas começam a ser reveladas. Os modelos de formação planetária sugerem que estes planetas de pequena massa, ao contrário dos planetas gigantes, devem ser frequentes em torno de estrelas de baixa metalicidade. Nesta tese, eu testo estas previsões estudando uma amostra de estrelas pobres em metais, em busca de planetas.

Nos próximos capítulos vou apresentar e desenvolver um método para a análise de dados de velocidade radial, com o qual conseguimos inferir simultaneamente os parâmetros orbitais dos planetas e também o número de planetas significativamente detectados num conjunto de dados. Também descreverei um critério de detecção rigoroso, baseado na quantidade de evidência que suporta a presença de sinais Keplerianos nos dados. Este método é depois estendido ao considerar processos Gaussianos como modelos para o ruído correlacionado presente nas variações de velocidade radial que são induzidas pela própria estrela. Estas variações, quando não tratadas, podem esconder verdadeiros sinais criados por planetas como a Terra. Ao modelar o ruído correlacionado, conseguimos também obter informação sobre a estrela e a actividade magnética à sua superfície.

A aplicação deste modelo ao sistema CoRoT-7 permitiu a detecção de dois planetas de baixa massa e levou também à medição do período de rotação estelar e da escala de tempo associada à evolução das regiões activas, usando apenas velocidades radiais. Com o mesmo método, não encontro evidência suficiente para a detecção dos planetas à volta das estrelas HD 175607 e HD 41248, que fazem parte da nossa amostra de estrelas pobres em metais.

Após a análise da amostra completa, com 109 estrelas, encontro uma frequência de planetas significativamente mais baixa do que os resultados para estrelas com metalicidade solar. Estes resultados representam um desafio para a teoria de formação planetária, e corroboram a ideia de que os planetas de pequena massa que se formam em discos de baixa metalicidade não migram até às regiões interiores do disco antes deste se dissipar, acabando com períodos orbitais mais longos que são, no presente, mais difíceis de detectar.

Contents

List of Figures

List of Tables

1

Introduction

Exoplanets are everywhere now. A field that did not exist 25 years ago is growing today at an exponential rate. In this introductory Chapter, I will briefly review the current state of the exoplanet field, summarizing some important discoveries from the last twenty-odd years. The tight connection between the study of exoplanets and of their host stars is also introduced here. My account is deliberately not exhaustive, instead aiming at setting the stage for the next chapters.

1.1 OVERVIEW OF THE EXOPLANET FIELD

What a daunting task it is to review a mere twenty years of exoplanet science! The number of planets keeps rising, almost on a daily basis (see Fig. 1.1), with more than 3500 now listed.[1] Astonishingly, about 80% of those planets have been discovered since I started working for this Thesis. But not only the numbers are impressive; also the diversity of these planets is quite exceptional, and completely unlike those of the Solar System.

Let us start with a brief historical account of exoplanet discoveries, grouped by the different detection methods. Some of the main results on planet demographics are then mentioned, with a special focus on planets detected with the radial velocity (RV) method. Finally I will mention some of the challenges posed by stellar activity.

[1] This number varies with the database one considers. The most commonly referenced are the "Extrasolar Planets Encyclopaedia" (exoplanet.eu – Schneider et al., 2011) and the "Exoplanet Orbit Database" (exoplanets.org – Wright et al., 2011; Han et al., 2014), even though others exist.

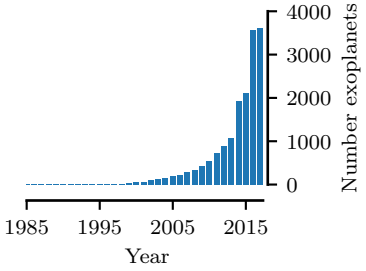


Figure 1.1: Cumulative number of exoplanets discovered over the years. Data from the exoplanet.eu database.

1.1.1 *Discoveries*

The story begins in the end of the 1980s, when the first tentative exoplanet detections were announced, even if somewhat cautiously. Campbell et al. (1988) discovered a 1.7 $M_{\rm Jup}$ (Jupiter masses) object orbiting γ Cep every 2.7 years, which would be later confirmed by Hatzes et al. (2003). In 1989, an 11 $M_{\rm Jup}$ companion to HD 114762 was found by Latham et al. (1989) and confirmed by Cochran et al. (1991) and Butler et al. (2006). Since the inclination angle of this system was unknown, the companion could easily be heavier than the deuterium burning limit ($\sim$13 $M_{\rm Jup}$; Spiegel et al. 2011), and therefore be either a brown dwarf or a very low mass star. A few years later, Hatzes and Cochran (1993) reported a 2.9 $M_{\rm Jup}$ companion to the K giant

HD 62509, which was later confirmed by Hatzes et al. (2006).

Around that time, the monitoring of radio pulsars led to the first convincing exoplanet discoveries around PSR1257+12 (Wolszczan and Frail, 1992) and later PSR1620-26 (Backer et al., 1993). These planets have very low masses, but are almost certainly completely different from the planets in the Solar System, in their properties, origin, and suitability for life (e.g. Agol, 2011; Martin et al., 2016).

But the title for the first exoplanet found around a solar-type star, and the discovery that truly changed (and kick-started) the field, goes to 51 Pegasi b, announced by Mayor and Queloz (1995). This planet is a *hot Jupiter*, with a minimum mass of 0.47 $M_{\rm Jup}$ and orbiting at 0.05 AU (Astronomical Units) from its host star every 4.2 days (Jupiter orbits at 5.2 AU from the Sun, with a period of about 11 years). Nothing in the Solar System had prepared us for this type of exoplanets (see e.g. Boss, 1995), but many of them quickly followed (e.g. Butler and Marcy, 1996; Marcy and Butler, 1996).

Exoplanets are commonly named after their host star, in alphabetical order starting with 'b' (Hessman et al., 2010).

Around one of the first known planet-host stars, Upsilon Andromedae, we would eventually find not just one but three orbiting planets (Butler et al., 1997, 1999): the first example of a multi-(exo)planet system. Around this time, RV surveys also started targeting M stars, allowing for the detection of a giant planet around the M4 dwarf Gl 876 (Marcy et al., 1998; Delfosse et al., 1998). Later works demonstrated that this system is actually quite complex, with two giant planets in a 2:1 mean motion resonance (Marcy et al., 2001) and a close-in super-Earth (Rivera et al., 2005).

In fact, the Gl 876 planetary system is not representative of the planet population later found around M dwarfs (e.g. Clanton and Gaudi, 2016).

Further improvements in RV precision and dedicated observing strategies, allowed for the detection of the first planets in the Neptune-mass range: μ Arae c (Santos et al., 2004a), GJ 436 b (Butler et al., 2004) and 55 Cnc e (McArthur et al., 2004). Since then, lower mass planets have been discovered (Fig. 1.2) and very low mass transiting planets (see below) were confirmed with follow-up RV observations.

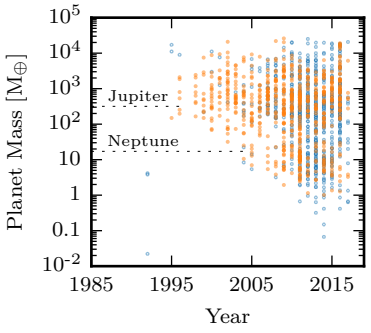


Figure 1.2: Planet masses over the years. Planets detected with the RV method are shown in orange, and in blue those detected with other methods.

No account of exoplanet discoveries would be complete without mentioning Proxima b, a planet with a minimum mass of only 1.3 $M_\oplus$, orbiting the Sun's closest stellar neighbor (Anglada-Escudé et al., 2016a). Moreover, at an orbital distance of around 0.05 AU, the planet is within Proxima Centauri's habitable zone, where water could be liquid on its surface (Ribas et al., 2016; Turbet et al., 2016).

* ◯ *

The first transiting planet was discovered near the end of the 20th century, with the detection of two transits of HD 209458 b by Charbonneau et al. (2000) and Henry et al. (2000). This detection was not only the first allowing for the derivation of the radius of an exoplanet (1.27 $R_{\rm Jup}$), but also a confirmation that these Jupiter-mass

planets in close orbits have radii and densities that are comparable to those of the gas giants of our own Solar System.

Several transit surveys started providing the first results around this time (e.g. Udalski et al., 2002b; Alonso et al., 2004; Holman et al., 2006). But it was not until more recently that we could finally start probing the population of planets with smaller radii (Fig. 1.3). This was largely due to the launch of the Kepler mission (Borucki et al., 2011), but the CoRoT mission (Barge et al., 2008) and ground-based surveys such as WASP (Pollacco et al., 2006), OGLE (e.g. Udalski et al., 2002a), HAT-Net (Bakos et al., 2007) or TrES (Alonso et al., 2004) played an important role on their own right.

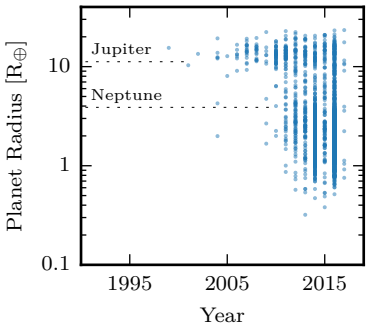


Figure 1.3: Planet radii over the years.

The transit method is particularly prone to astrophysical false-positive scenarios, created by configurations of two or more stars which may not involve an exoplanet (see, e.g., Cameron, 2012; Santerne et al., 2013). This means that the RV follow-up of candidate planet signals or their statistical validation is usually required to confirm a transiting planet detection. A classic example of a false positive is the case of OGLE-TR-33 (Torres et al., 2004), a planet candidate later found to be consistent with a hierarchical triple system composed of an F6 star and an M0-F4 eclipsing binary. Some of the currently known exoplanets with smallest radii and lowest masses were first discovered in transits and followed up with high-precision RV spectrographs for confirmation (e.g. Queloz et al., 2009; Pepe et al., 2013; Dumusque et al., 2014b; Marcy et al., 2014; López-Morales et al., 2016).

In some cases, however, transiting planets in multi-planet systems have non-Keplerian orbits, causing the times and durations of the transits to vary. The analysis of these transit timing variations (TTV) and transit duration variations (TDV) allows for the characterization of the planets' masses and sometimes the detection of additional planets (e.g. Holman and Murray, 2005) avoiding the need for follow-up RV observations. The first convincing detection of a TTV signal happened for the system of two Saturn-like planets orbiting Kepler-9 (Holman et al., 2010), and since then a large number of systems have been discovered showing both TTVs and TDVs (e.g. Carter et al., 2012; Holczer et al., 2016).

Measurements of both the planet's radius and mass yield an estimate of its mean density, informing interior composition models (e.g. Wolfgang and Lopez, 2015, and references therein). This allows, in some cases, to infer whether a given planet might have a rocky or gaseous interior (e.g. Weiss and Marcy, 2014; Dorn et al., 2017).

* ◯ *

The first exoplanet to ever pose for a picture was 2M1207-39 b,

announced by Chauvin et al. (2004), while other high-profile detections followed a few years later (Fomalhaut b; Kalas et al., 2008, HR 8799 b,c,d; Marois et al., 2008 and β Pictoris b; Lagrange et al., 2010). The technique of direct imaging is one of the few direct detection methods, observing light from the planets themselves. Giant planets on long-period orbits are much easier to image, even if this still requires considerable observational effort (e.g. Claudi, 2016).

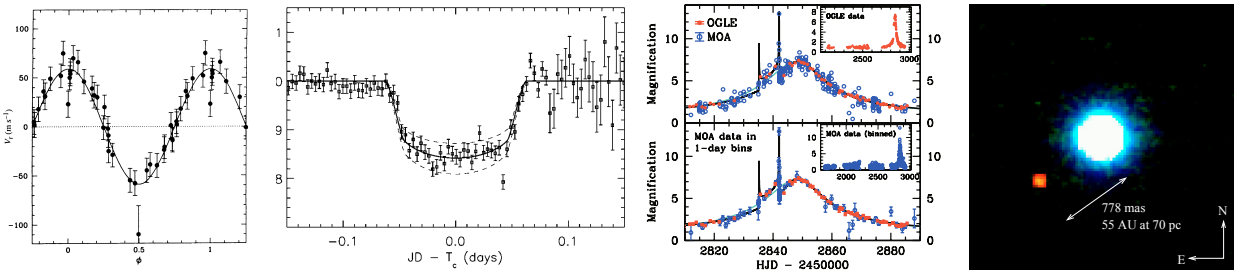


Figure 1.4: A collection of firsts. From left to right, the first exoplanets detected with the RV method (51 Peg b; Mayor and Queloz, 1995), the transit method (HD 209458 b; Charbonneau et al., 2000; Henry et al., 2000), gravitational microlensing (OGLE-2003-BLG-235L b; Bond et al., 2004) and direct imaging (2M1207-39 b; Chauvin et al., 2004). Figures adapted from the respective discovery papers.

High-contrast adaptive optics instruments, such as SPHERE (Beuzit et al., 2008) and GPI (Macintosh et al., 2008), are now probing the population of exoplanets with orbital separations beyond 10 AU and masses larger than 1 $M_{\rm Jup}$, calling for new theories of planet formation and migration (see e.g. Bowler, 2016). Parallel to the detection of these populations, direct imaging provides exquisite information about the chemical compositions, internal structures, atmospheric dynamics, and overall physical properties of planets (e.g. Bonnefoy et al., 2016).

In its turn, the technique of gravitational microlensing is arguably more exotic, as it does not rely on the detection of light from either the host star or the planet: the planets are instead discovered by their gravitational perturbation to the light of a more distant source (see e.g. Gould, 2016). This comes with a few idiosyncrasies. The technique is sensitive to very low mass planets on wide orbits and also to free-floating planets. But microlensing events are rare, they cannot be repeated or predicted, and seldom provide detailed information about the host star (e.g. Park et al., 2006; Gaudi, 2012). Follow-up spectroscopic observations allow for a more precise characterization of some microlensing events (e.g. Santerne et al., 2016a).

The first microlensing planet, OGLE-2003-BLG-235L b, announced by Bond et al. (2004), is probably[2] a 1.5 $M_{\rm Jup}$ planet orbiting at about 3 AU. Since then, the number of microlensing detections has grown, revealing interesting new populations such as cold Neptunes (Beaulieu et al., 2006; Bennett et al., 2008; Muraki et al., 2011), which turn out to be the most common type of planets beyond the snow line (Suzuki et al., 2016), and free-floating planets (Sumi et al., 2011), as well as evidence for at least one planet for every star in the Milky Way (Cassan et al., 2012).

[2] Conditioned on the lens being a main-sequence star.

* ◯ *

The search for exoplanets with astrometry also merits a mention, even if no detections have been made so far. From very early on, claims of planet detections started long-running debates. For examples, the claimed detections of a 10 $M_{\rm Jup}$ companion around 70 Oph by Reuyl and Holmberg (1943), and a 16 $M_{\rm Jup}$ companion to 61 Cyg by Strand (1943) were only disproved by Heintz (1978). As of today, the only existing detections are of planets or brown-dwarfs that were first detected using the RV technique (e.g. Benedict et al., 2006).

The predicted astrometric variations of the known exoplanets are typically below the 1 milliarcsecond accomplished by state-of-the-art instruments (e.g. Santos et al., 2016). But this situation will change with the exquisite astrometric precision of the Gaia mission (Perryman et al., 2001), which promises to find a large number of giant planets. Recent simulations predict more than 21 000 long-period planets (with masses around $1-15$ $M_{\rm Jup}$) should be discovered by the 5 year nominal mission (Perryman et al., 2014 see also Sozzetti et al., 2001), which will certainly add to our understanding of the giant planet population.

1.1.2 *Demographics*

Natural follow-up questions to whether or not other stars have planets, may be "how many stars have planets?" and "what kind of planets?". Here, one of the main goals is to determine the frequency of stars with planets, given simply by

$$f_p = \frac{\text{number of stars with planets}}{\text{total number of stars}} \tag{1.1}$$

although any given planet survey has its own biases and selection effects, which must obviously be taken into account. The frequency of planets can be calculated for a given population of host stars and for planets with different orbital characteristics.

One might also be interested in the planet occurrence rate, which also takes into account multi-planet systems. High multiplicity rates within a survey imply that the fraction of stars hosting a planetary system (f_p) can be much smaller than the number of planets per star. The latter relates to the occurrence rate directly (see Youdin, 2011).

Besides planet occurrence, we can also study exoplanet *populations*, as well as the relative numbers of planets with different masses, radii, orbital distances, periods, and eccentricities. Fig. 1.5 shows the masses and orbital periods of all the planets where those parameters were measured (a total of 3550 planets, up to March 2017). Three distinct clusters are visible in this particular parameter space.

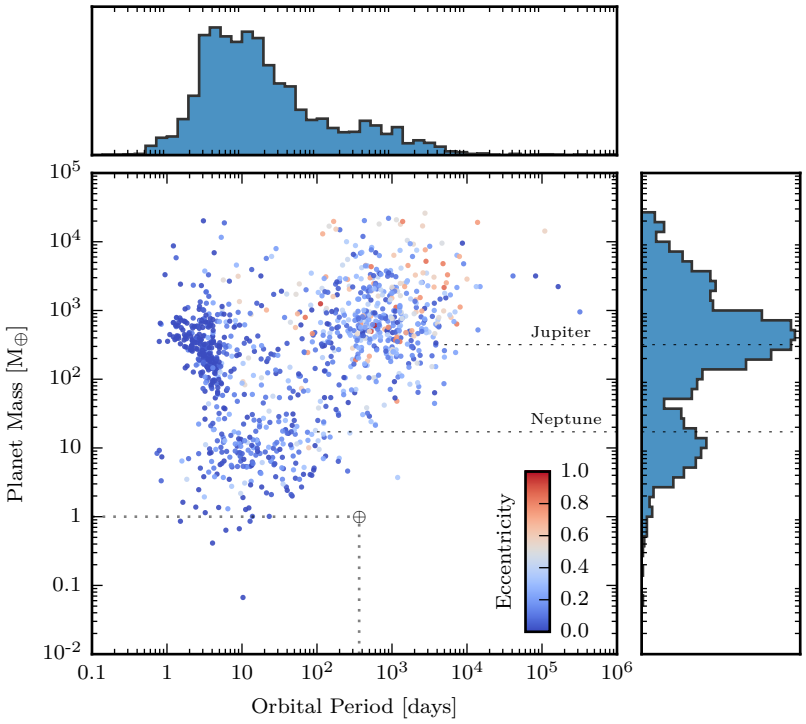


Figure 1.5: Distribution of masses and orbital periods for all planets whose mass has been measured. The colour scale in the centre panel represents the orbital eccentricity. Note that the masses of planets detected with the RV method are minimum masses. Data obtained from the `exoplanet.eu` database.

Focusing now on the results of RV surveys (e.g. Udry and Santos, 2007), we can start by looking at hot-Jupiters, for instance. These Jupiter-mass planets, in circular orbits and orbital periods less than 10 days, are now known to be relatively rare, with an occurrence rate close to 1% (Marcy et al., 2005; Mayor et al., 2011; Wright et al., 2012). But the fact that they exist and their orbital properties provide important clues to their formation process (e.g. Batygin et al., 2016, and references therein).

Because of their large masses, hot-Jupiters must have a substantial gaseous component and must have formed before the dissipation of the proto-planetary disk. Therefore, in situ formation, although possible, is unlikely (e.g. Bodenheimer et al., 2000). The formation of a solid core at larger separations from the host star, followed by accretion of a gaseous envelope and migration to their current location is the preferred explanation (e.g. Alibert et al., 2005, but see also Boley et al., 2016).

On a different region of parameter space, RV surveys also revealed a population of gas giants at larger orbital distances, between 1 and 5 AU (see Fig. 1.5) having an overall occurrence rate of about 15% (see e.g. Udry and Santos, 2007, and references therein).

The mass distribution of these giant planets peaks around 3 $M_{\rm Jup}$ and presents a long tail toward masses larger than 10 $M_{\rm Jup}$. Within the brown-dwarf regime (masses between $\sim$ 15 $M_{\rm Jup}$ and $\sim$ 60 $M_{\rm Jup}$) the number of detections is very small, in what has been called the

"brown-dwarf desert" (e.g. Grether and Lineweaver, 2006; Grieves et al., 2017). This population of giant planets also shows a wide distribution of orbital eccentricities, unlike the low eccentricities seen in the Solar System. The giant planet eccentricity distribution, with a median value of 0.29, is closer to that of binary stars (see Udry and Santos, 2007, their Figure 6).

Many of these giant planets are also found in systems (unlike hot-Jupiters), with various dynamical configurations (e.g. Correia et al., 2009; Wright et al., 2009). In some multi-planet systems, the gravitational interactions between planets are strong enough to be detectable in RV measurements[3], allowing for the orbital inclination angles and the true masses of the planets to be measured (e.g. Correia et al., 2010). These systems are interesting for benchmarking theories of planet formation and migration (e.g. Ford, 2014).

[3] See Section 2.1.4.

Since the detection of low-mass planets requires a large observational effort, only a few surveys gathered enough detections to allow for statistical studies. Results from the HARPS survey of FGK stars (Mayor et al., 2011), the HARPS survey of M dwarfs (Bonfils et al., 2013), and the Keck-HIRES survey of FGK stars (Howard et al., 2010), suggest a large population of Neptunes and super-Earths in short period orbits. The occurrence rate of these planets, with $M \sin i < 30\, M_\oplus$ and $P < 50$ days is about 30% around FGK stars and 40% in M dwarfs. Most low-mass planets are also found in multi-planet systems (e.g. Lovis et al., 2011b)

The high occurrence rates tell us that systems of multiple planets with masses between 1 $M_\oplus$ and 20 $M_\oplus$, orbiting within $0.5 - 1$ AU, might actually be the most common planetary systems in the Galaxy. This remarkable result has been confirmed with transit surveys (see e.g. Lissauer et al., 2014) and is in agreement with planet population synthesis models based on the core-accretion theory (Ida and Lin, 2004; Mordasini et al., 2012, 2015).

1.1.3 *Limitations and challenges*

Any exoplanet detection method has instrumental limitations, which determine the properties of the detectable planets. Often, these limitations can be overcome or improved with technological advances that increase the instrumental precision and accuracy. This was the case in the past for the transit and RV methods, with Kepler and HARPS, and will be so in the future for the microlensing and astrometry methods, with WFIRST and GAIA, just to give some examples.

Adding to the technical challenges, the detection and characterization of other planets also has limitations that originate in the unavoidable study of the planet-host stars. This stellar "noise" is created by different processes, with the phenomena related to stellar

activity, stellar granulation, and oscillations being particularly nasty for exoplanet hunters. They can prevent us from finding planets, if the perturbation is of comparable amplitude or larger than the orbital RV variation, or even introduce false candidates, if they produce a periodic and stable signal over a few stellar rotations (e.g. Forveille et al., 2009; Figueira et al., 2010).

Furthermore, these physical phenomena produce signals with different timescales[4], from several years, related with long term magnetic cycles (e.g. Santos et al., 2010b; Lovis et al., 2011a) down to a few minutes, from the stellar oscillation modes (e.g. Dumusque et al., 2011c). On the shorter timescales, the effects of oscillations and granulation can often be circumvented with dedicated observing strategies (Dumusque et al., 2011b,c). On longer timescales, close to the rotation period of the star, the signals created by magnetic fields in the stellar surface (e.g. spots, faculae, convective changes), strongly affect both the photometric and RV signals (e.g. Meunier et al., 2010b).

[4] See for example Fig. 3.2.

In the particular case of the RV method, measurements of different activity indicators (see section 3.2.2) are often used as diagnostics of activity effects. However, recent examples show that the relation between these indicators and stellar activity is not always completely understood, leading to signals that almost perfectly mimic the signature of planets (Jenkins and Tuomi, 2014; Robertson and Mahadevan, 2014; Santos et al., 2014; Anglada-Escudé and Tuomi, 2015).

For transit surveys, stellar activity is also particularly relevant. Not only can it induce strong out-of-transit photometric modulations that need to be filtered or modeled (e.g. Aigrain et al., 2015; Barclay et al., 2015), but also in-transit fluctuations that hinder the precise measurement of the transit depth, and hence the planet radius (e.g. Oshagh et al., 2013).

Different sources of noise represent a big challenge in planet detection and characterization. The understanding of the different physical phenomena as well as ways to model or subtract activity-induced signals is now one of the most important avenues to guarantee the success of future ground and space based exoplanet projects. In Chapter 3, I will present a method for the analysis of RV data that can deal with strong activity contaminations and still allow for the detection of the planetary signals, taking us one step further in the right direction.

1.2 STAR-PLANET CONNECTION

The study of planet host stars is, in many ways, fundamental to the understanding of the properties and formation mechanisms of exoplanets. On one hand, the determination of planet properties is often

limited by the imperfect knowledge of their host stars, with stellar mass and radius being the usual sources of uncertainties (e.g. Torres et al., 2008; Marcy et al., 2014). On the other hand, the chemical compositions of the planet, the protostellar disk, and the stellar atmosphere are intimately linked through the planet formation process. Therefore, precise chemical abundances of host stars provide important clues in understanding the planets and their observed properties (e.g. Adibekyan et al., 2012b; Santos et al., 2015).

The planet-metallicity correlation is one example of this connection: it is one of the most striking and best established results for the population of giant planets that their host stars have a higher average metallicity, when compared with field stars (e.g. Gonzalez, 1997; Santos et al., 2001, 2004b; Fischer and Valenti, 2005).

At first, this correlation was thought to be the result of pollution during the planet formation process. If short-period Jupiters were formed outside their current orbits, and then migrated inwards, most of the disc material between the planet and the star would be accreted onto the star, increasing the envelope metallicity significantly. Later works found no evidence for this explanation, suggesting instead that high metallicity in planet-host stars is inherited from the primordial cloud (Santos et al., 2001; Fischer and Valenti, 2005).

The clear correlation between the presence of giant planets and metallicity is visible in the top panel of Fig. 1.6, which shows the giant planets from the HARPS survey (Mayor et al., 2014). The complete distribution may not be well described by a single power-law (as suggested for the metal-rich regime by Johnson et al., 2010), and may actually be flat for metallicities below -0.1 dex, with a lower limit of -0.7 dex, below which no giant planets might be able to form (Mortier et al., 2013; Hasegawa and Hirashita, 2014).

This correlation is linked to the planet formation process, and its existence is generally accepted to be in line with core-accretion being the main process of formation of giant planets (e.g. Mordasini et al., 2012; Hasegawa and Pudritz, 2014). In very basic terms, giant planet formation by core accretion takes place in two steps: first, the planetary cores form by runaway and oligarchic growth (e.g. Kokubo and Ida, 2002) after which gas starts accreting onto the cores (Pollack et al., 1996; Ida and Lin, 2004) if the mass of the planetary core grows to exceed a critical mass (of around 10 $M_\oplus$; Lissauer et al. 2009). In proto-planetary disks with higher metallicities, the cores form faster and in time for accretion to start before the disk dissipates. In lower-metallicity disks, the cores do not grow fast enough to accrete gas in large quantities before disk dissipation, resulting in a lower fraction of giant planets (Mordasini et al., 2009b).

But it is important to mention that the disk instability model, an

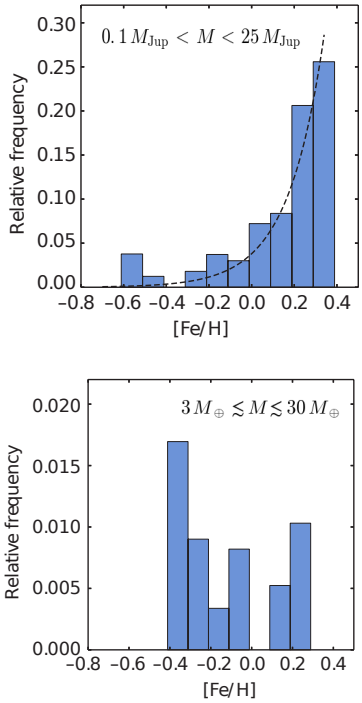


Figure 1.6: Planet occurrence as a function of metallicity. There is a clear correlation between the presence of giant planets and the metallicity of the star, but this trend is not present for stars hosting lower-mass planets (Mayor et al., 2011; Sousa et al., 2011b). Adapted from Mayor et al. (2014).

alternative to core-accretion, can also explain the metallicity-giant-planet correlation, within its more recent incarnations (e.g. Nayakshin, 2017). This theory also predicts a change in the metallicity correlation to happen around the mass of $\sim 5\ M_{\rm Jup}$, which is in line with recent observations (Santos et al., 2017).

* ◯ *

For planets with lower masses, the metallicity dependence of the occurrence rate is clearly weaker if it exists at all (Fig. 1.6). These planets seem to be found around stars with a large range of metallicities, with both RV (Mayor et al., 2011; Sousa et al., 2011b) and transit (Buchhave et al., 2012; Buchhave and Latham, 2015) surveys confirming this result (for planets with low masses and small radii, respectively). Note, however, that a universal metallicity dependence for all planet sizes has also been argued for (Wang and Fischer, 2014), and Zhu et al. (2016) showed that neither hypothesis can be ruled out with currently available data.

Part of the difficulty in studying this issue using the planet population from Kepler is the determination of accurate metallicities for such faint targets (e.g. Dong et al., 2014). Another problem is related to the construction of a "control" sample of stars known *not* to host planets, with which the metallicity distribution of planet-hosts can be compared. RV surveys have some advantage in this regard: the spectroscopic observations allow for more accurate metallicity determinations and it is fairly straightforward to rule out the presence of planets down to a certain mass (see section 4.4.1).

More recent works have used only the exoplanets with measured masses and seem to suggest differences in the metallicity dependence for Neptune-like planets (with masses between 10 and 40 $M_\oplus$) and super-Earths ($< 10\ M_\oplus$). Courcol et al. (2016) analysed 157 low-mass exoplanets and suggested the existence of an upper boundary of planet mass that increases with metallicity and also depends on the orbital period (see Fig. 1.7).

These results are again in general agreement with core-accretion theory: planets around metal-poor stars would start with smaller cores, and thus accrete gas less efficiently (Dawson et al., 2015). In this scenario, the more metal-poor the disc (and therefore the star) the less efficient it is to form Neptune-like planets, with rocky cores and icy or gaseous outer layers.

One interesting hypothesis to explain the dependence between the mass boundary and the orbital period, as found by Courcol et al. (2016), is that planets would actually form at all metallicities (Hasegawa and Pudritz, 2014), but in metal-poor discs Neptune-like planets would form at larger distances from the star, or would not

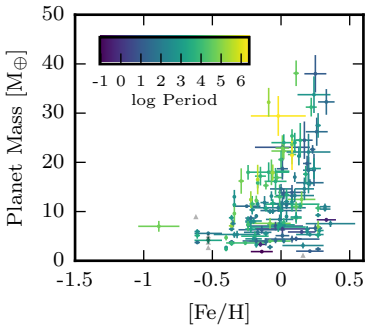


Figure 1.7: Mass-metallicity distribution of low-mass planets, with the logarithm of the orbital period as the colour scale. The same data as in Courcol et al. (2016) are shown here, except that α Cen B b, HD 175607 b and GJ 667C d, e, f and g have been masked as triangles because of doubts about their planetary nature (see Robertson and Mahadevan, 2014; Rajpaul et al., 2016, and section 4.3.3).

migrate as quickly as in metal-rich discs (see for example Ercolano and Clarke, 2010). These planets would therefore still be undetectable with current RV surveys.

Earlier studies had already hinted at this relation between metallicity and the planets' orbital period. Adibekyan et al. (2013) analysed a large sample of FGK planet-hosts and found that planets orbiting metal-poor stars have longer orbital periods than those in metal-rich systems. Similar conclusions were reached by Beaugé and Nesvorný (2013), who found a lack of super-Earths with small orbital periods around metal-poor stars. These results also support the idea that smaller planets form farther from their host star and do not migrate as far as the giants.

From the theoretical point of view, the study of small and low-mass planets is less developed than that of giant planets. The models of Mordasini et al. (2009a, 2012) predict the absence of a metallicity correlation for Neptune-mass planets and an inverse metallicity effect for low-mass planets. But these models are, first and foremost, giant-planet formation models so that the behaviour of the lower mass population is still open to debate.

The study of specific elemental abundances also gives further insight into the planet-formation process. The abundance of α-elements, for example, plays an important role in the formation of planetary systems, specially in metal-poor environments (Adibekyan et al., 2012a). Abundances of other chemical elements, such as lithium (Reddy et al., 2002; Israelian et al., 2009; Baumann et al., 2010) and refractory elements, have also been identified as possible signatures of planet engulfment or terrestrial planet formation (González Hernández et al., 2010; Ramírez et al., 2010). Another example of the chemical connection between stellar hosts and planets is the recent work of Santos et al. (2015) who used precise stellar chemical abundances to constraint the iron mass fraction of the rocky planets CoRoT-7, b Kepler-10 b, and Kepler-93 b (see also Dorn et al., 2015; Unterborn et al., 2017, for example).

1.3 THIS THESIS

It is hopefully clear from the above paragraphs that the population of low-mass planets orbiting low-metallicity stars is still not completely understood. Whether the planet-metallicity correlation extends to lower mass planets is still an open question, and exploring its possible answers can provide insights into planet formation. Current observational results are in good agreement with the theory of core-accretion, but some key physical processes in that theory can be tested by studying the properties of exoplanets around metal-poor stars.

In this context, I have identified the main scientific driver for my Thesis: assessing the frequency of low-mass planets around metal-poor stars, using the exquisite data from a RV survey specifically targeted at these objects. In order to reach this goal, I first needed to explore and develop new methods for detecting exoplanets in RV data of active and non-active stars.

This manuscript is organized as follows. In Chapter 2, I introduce the RV method in detail, explain how we measure the RV of stars, and describe a model for the analysis of RV datasets, based on the work of Brewer and Donovan (2015). This model is then extended, in Chapter 3, with a stellar activity component, and applied to RV data of CoRoT-7, following the analysis first published in Faria et al. (2016a). Chapter 4 deals with a sample of about one hundred metal-poor stars, presenting the analysis of the RV timeseries and the calculation of the occurrence rate of low-mass planets in this metallicity regime. A preliminary analysis of this sample was first published as Faria et al. (2016b).

By the end of the Thesis, in Chapter 5, I will present the main conclusions of my work: a general model for the detection of exoplanet signals in RV data, which can deal with stellar activity contaminations, and the lack of low-mass planets in short-period orbits around metal-poor stars. These two results set the stage and open roads for future exploration, which I also discuss.

2

Planet detection in radial velocities

"Truth is ever to be found in the simplicity, and not in the multiplicity and confusion of things."
— Isaac Newton

In 1952, Otto Struve writes:

> But how should we proceed to detect [objects of planet-like character]? (...) It is not unreasonable that a planet might exist at a distance of 1/50 astronomical unit (...) Its period around a star of solar mass would then be about 1 day (...) If the mass of this planet were equal to that of Jupiter, it would cause the observed radial velocity of the parent star to oscillate with a range of $\pm 0.2\,\mathrm{km \cdot s^{-1}}$. (Struve, 1952)

almost anticipating[1] the moment when, forty years later, the radial-velocity (RV) observations of 51 Peg, made with the ELODIE spectrograph in France, show the star moving back and forth with a period of 4.23 days and a semi-amplitude of 56 $\mathrm{m \cdot s^{-1}}$ (Mayor and Queloz, 1995). The RV method has since been incredibly successful, detecting hundreds of exoplanets orbiting hundreds of stars.

[1] See also Belorizky (1938).

In this Thesis, I will use the RV method to search for low-mass planets around low-metallicity stars. This first chapter serves to set the theoretical grounds for how these detections are made.

Starting with a derivation of the RV signal caused by an orbiting planet (Section 2.1), I then proceed with a brief explanation of how high-resolution spectroscopy is used to measure the radial velocities of stars with $\mathrm{m \cdot s^{-1}}$ precision (Section 2.2). We then venture into the analysis of RV datasets (Section 2.3), building a custom probabilistic model for the problem, which we apply in Section 2.4.

2.1 THE KEPLERIAN CURVE

When a planet orbits its host star, both bodies move around their common center of mass. It is a remarkable (albeit long) exercise in geometry and calculus to derive the expression for the velocity of the star in the direction of an observer. Because this expression forms one of the theoretical cores of the radial velocity method, I present a complete and self-consistent derivation below.

This derivation (or part of it) is covered in a number of texts on celestial mechanics (e.g. Moulton, 1970), and follows closely the treatment of Murray and Correia (2010). The reader may feel free to skip to Section 2.1.3, after a quick look at Fig. 2.3 for a visual meaning of the orbital parameters.

2.1.1 *The two body problem*

Assume a system of two point-mass bodies with masses m_1 (the star) and m_2 (the planet). They will attract each other with a force proportional to the product of their masses and inversely proportional to the square of the distance r between their centers. This is a statement of Newton's universal law of gravitation:

$$F = \mathcal{G}\,\frac{m_1\,m_2}{r^2} \tag{2.1}$$

where $\mathcal{G}$ is the universal gravitational constant.

$\mathcal{G} = 6.674 \times 10^{-8}\,\mathrm{cm^3 g^{-1} s^{-2}}$

Let the star and planet be positioned relative to an origin O, with position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, as in Fig. 2.1. The relative motion of the planet with respect to the star is given by the vector $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and the gravitational forces acting on the star and the planet are, respectively,

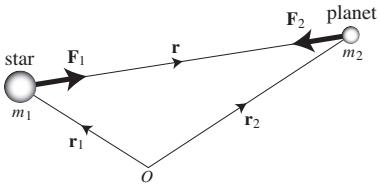


Figure 2.1: Forces acting on a star with mass m_1 and a planet with mass m_2. The bodies are positioned relative to the origin O. From Murray and Correia (2010).

$$\mathbf{F}_1 = m_1\ddot{\mathbf{r}}_1 = +\mathcal{G}\,\frac{m_1 m_2}{r^3}\,\mathbf{r}\,, \tag{2.2}$$

$$\mathbf{F}_2 = m_2\ddot{\mathbf{r}}_2 = -\mathcal{G}\,\frac{m_1 m_2}{r^3}\,\mathbf{r}\,. \tag{2.3}$$

We start by considering the motion of the planet relative to the star. With $\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_2 - \ddot{\mathbf{r}}_1$ and Eqs. (2.2) and (2.3), we obtain the equation of relative motion,

$$\ddot{\mathbf{r}} + \mathcal{G}(m_1 + m_2)\frac{\mathbf{r}}{r^3} = 0\,. \tag{2.4}$$

Taking the vector product of $\mathbf{r}$ with Eq. (2.4) we have $\mathbf{r} \times \ddot{\mathbf{r}} = 0$, which after integration yields

$$\mathbf{r} \times \dot{\mathbf{r}} = \mathbf{h} \tag{2.5}$$

where $\mathbf{h}$ is a constant vector, perpendicular to both $\mathbf{r}$ and $\dot{\mathbf{r}}$. This implies that the motion of the planet around the star occurs in a plane (the *orbital plane*) perpendicular to the direction defined by $\mathbf{h}$, and that the position and velocity vectors will always lie in the same plane.

We now change to a polar coordinate system (r, θ) centered on the star. The position, velocity and acceleration vectors are written as

$$\mathbf{r} = r\,\hat{\mathbf{r}} \tag{2.6}$$

$$\dot{\mathbf{r}} = \dot{r}\,\hat{\mathbf{r}} + r\dot{\theta}\,\hat{\boldsymbol{\theta}} \tag{2.7}$$

$$\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\theta}^2)\hat{\mathbf{r}} + \left[\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}\left(r^2\dot{\theta}\right)\right]\hat{\boldsymbol{\theta}}\,. \tag{2.8}$$

where $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are unit vectors along the coordinate axis. Substituting Eq. (2.7) into Eq. (2.5) gives us the magnitude of the vector $\mathbf{h}$:

$$h = r^2\dot{\theta}\,. \tag{2.9}$$

The area dA swept by the star-planet vector in the time interval dt is

$$dA = \int_0^r r\,dr\,d\theta = \frac{1}{2}r^2 d\theta\,, \tag{2.10}$$

and thus

$$\dot{A} = \frac{1}{2} r^2 \dot{\theta} = \frac{1}{2} h = \text{constant} \tag{2.11}$$

which is equivalent to Kepler's second law of planetary motion, stating that the star-planet line sweeps out equal areas in equal times.

"*Arcum ellipseos, cujus moras metitur area AKN, debere terminari in LK, ut sit AM.*"

— Kepler (1609, pp. 294)

Equating the radial components of Eqs. (2.4) and (2.8) gives the scalar differential equation

$$\ddot{r} - r\dot{\theta}^2 = -\frac{\mathcal{G}(m_1 + m_2)}{r^2} \,. \tag{2.12}$$

To find r as a function of θ, we make the substitution $u = 1/r$. Differentiating r with respect to time and using Eq. (2.9) we can eliminate time in the differential equation. This gives

$$\ddot{r} = -h \frac{\mathrm{d}^2 u}{\mathrm{d}\theta^2} \dot{\theta} = -h^2 u^2 \frac{\mathrm{d}^2 u}{\mathrm{d}\theta^2} \tag{2.13}$$

and hence Eq. (2.12) can be written

$$\frac{\mathrm{d}^2 u}{\mathrm{d}\theta^2} + u = \frac{\mathcal{G}(m_1 + m_2)}{h^2} \,. \tag{2.14}$$

Binet's equation

which is a second order, linear differential equation, with the general solution

$$u = \frac{\mathcal{G}(m_1 + m_2)}{h^2} \left[1 + e\cos(\theta - \varpi)\right] \,, \tag{2.15}$$

where e and ϖ are two constants of integration. Substituting back for r gives

$$r = \frac{p}{1 + e\cos(\theta - \varpi)} \,, \tag{2.16}$$

with $p = h^2/\mathcal{G}(m_1 + m_2)$. Equation (2.16) is the general equation of a set of curves known as conic sections where e is the *eccentricity* and p is a constant called the *semilatus rectum*.

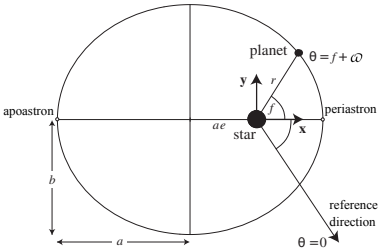


Figure 2.2: Geometry of the elliptical orbit of the planet around the star. Adapted from Murray and Correia (2010).

The initial conditions determine which particular conic section (a circle, ellipse, parabola, or hyperbola) the planet follows. We consider only circular and elliptical motions, for which

$$0 \leq e < 1 \qquad\qquad p = a(1 - e^2) \,, \tag{2.17}$$

where a is the *semi-major axis* of the ellipse (or the radius of the circle). The semi-major axis and the eccentricity are geometrically related by

$$b^2 = a^2(1 - e^2) \,, \tag{2.18}$$

where b is the *semi-minor axis* of the ellipse (see Fig. 2.2).

Therefore, for a given position in the orbit, corresponding to a given value of θ, the distance between the planet and the star is

$$r = \frac{a(1 - e^2)}{1 + e\cos(\theta - \varpi)} \,. \tag{2.19}$$

The path of the planet around the star is an ellipse with the star at one focus: this is Kepler's first law of planetary motion.

"Ergo ellipsis est Planetæ iter"
— Kepler (1609, pp. 285)

The angle θ is called the true longitude. From Eq. (2.19), the minimum and maximum values of r are $a(1-e)$ at $\theta = \varpi$ and $a(1+e)$ at $\theta = \varpi + \pi$, respectively. These points are called the *periastron* and the *apastron* of the orbit.

The angle ϖ is called the *longitude of periastron* of the planet's orbit and gives the angular location of the closest approach with respect to the reference direction $\theta = 0$. Defining the *true anomaly* to be the angle $f = \theta - \varpi$ (see Fig. 2.2), then Eq. (2.19) can be written

$$r = \frac{a(1-e^2)}{1+e\cos f}\,. \tag{2.20}$$

where f is measured with respect to the periastron direction.

Even with no reference to time in the equation of motion, we can relate the orbital period P to the semi-major axis, a. The area of an ellipse is $A = \pi ab$ and this is swept out by the star-planet line in a time P. Therefore, from Eq. (2.11), $A = hP/2$ and so

$$P^2 = \frac{4\pi^2}{\mathcal{G}(m_1+m_2)}\,a^3\,, \tag{2.21}$$

which is Kepler's third law of planetary motion. It implies that the period of the planet's orbit is only a function of the sum of the masses of the star and planet, and the semi-major axis.

"Sed res est certissima exactissimaque quod proportio qua est inter binorum quorumcunque Planetarum tempora periodica, sit præcise sesquialtera proportionis mediarum distantiarum"
— Kepler (1619, pp. 294)

Since ϖ is a constant, $\dot\theta = \dot f$ from the definition of the true anomaly, and Eq. (2.7) gives

$$v^2 = \dot{\mathbf{r}}\cdot\dot{\mathbf{r}} = \dot r^2 + r^2\dot f^2\,. \tag{2.22}$$

By differentiating Eq. (2.20) we obtain

$$\dot r = \frac{r\dot f\, e\sin f}{1+e\cos f}\,. \tag{2.23}$$

If we now define the mean motion of the planet's orbit

$$n = \frac{2\pi}{P} \tag{2.24}$$

and use Eqs. (2.9) and (2.16), we can write

$$\dot r = \frac{na}{\sqrt{1-e^2}}\,e\sin f \tag{2.25}$$

and

$$r\dot f = \frac{na}{\sqrt{1-e^2}}(1+e\cos f)\,. \tag{2.26}$$

Therefore we can rewrite Eq. (2.22) as

$$v^2 = \frac{n^2a^2}{1-e^2}(1+2\,e\cos f+e^2)\,. \tag{2.27}$$

or, after some manipulation,

$$v^2 = G(m_1 + m_2)\left(\frac{2}{r} - \frac{1}{a}\right) \tag{2.28}$$

which shows the dependence of v on r.

At this stage we have solved the equation of motion of the two-body problem to obtain the path of the planet with respect to the star. But up to now we are only able to calculate r as a function of θ, not as a function of time. To do the latter, we first derive an expression for $\dot{r}$ in terms of r. We can do this by using Eqs. (2.20), (2.26) and (2.28) to rewrite Eq. (2.22) as

$$\dot{r}^2 = n^2 a^3 \left(\frac{2}{r} - \frac{1}{a}\right) - \frac{n^2 a^4 (1 - e^2)}{r^2}\,. \tag{2.29}$$

which simplifies to

$$\dot{r} = \frac{na}{r}\sqrt{a^2 e^2 - (r - a)^2}\,. \tag{2.30}$$

In order to solve this differential equation we introduce a new variable, E, the *eccentric anomaly*, by substituting

$$r = a(1 - e\cos E)\,. \tag{2.31}$$

The equation transforms to

$$\dot{E} = \frac{n}{1 - e\cos E}\,. \tag{2.32}$$

and its solution can be written as

$$n(t - t_0) = E - e\sin E\,, \tag{2.33}$$

where t_0, called the *time of periastron passage*, is the constant of integration and we used the boundary condition $E = 0$ when $t = t_0$. At this point we can define a new quantity, M, the *mean anomaly* such that

$$M = n(t - t_0)\,, \tag{2.34}$$

Note that $M = f = 0$ when $t = t_0$ or $t = t_0 + P$ (passages at periastron) and $M = f = \pi$ when $t = t_0 + P/2$ (apoastron passage). We can further write

$$M = E - e\sin E\,. \tag{2.35}$$

Kepler's equation

The solution of this equation is the fundamental step to find the orbital position at a given time. For a particular time t, we find M from Eq. (2.34), then E by solving Kepler's equation, Eq. (2.35), and finally find r using Eq. (2.31), and f using Eq. (2.20).

Solving the transcendental equation (2.35) is an interesting mathematical and computational task. In the numerical implementations developed here, I used the method outlined by Murison (2006).

2.1.2 *The three-dimensional orbit*

So far we have shown that the orbit of the planet is confined to a plane and derived the position in the orbit as measured from an arbitrary reference direction. We will now "place" the orbital plane in the three-dimensional coordinate system of the observer.

Consider the position vector of the planet,

$$\mathbf{r} = x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}} + 0\,\hat{\mathbf{z}}\,. \tag{2.36}$$

The x and y axes lie in the orbital planet and the z-axis is perpendicular to it (because the orbit is constrained to the orbital plane, the z coordinate is zero). The x-axis, by definition, lies along the major axis of the ellipse, in the direction of periastron.

Now consider a second coordinate system with axes (X, Y, Z). We will take the plane (X, Y) to be the plane of the sky, perpendicular to the line of sight and the Z-axis to be oriented towards the observer.

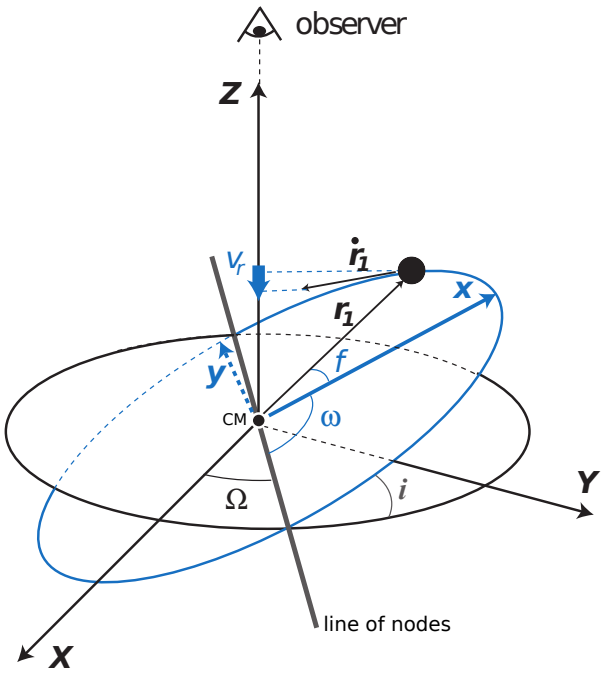


Figure 2.3: Representation of the orbital and sky-observer coordinate systems. The stars' velocity $\dot{r}_1$ around the center of mass (CM) has a component along the line of sight which corresponds to the radial velocity v_r. Three angles provide the transformation between the two coordinate systems: the longitude of the ascending node Ω, the inclination i and the argument of periastron ω. The true anomaly f is also shown. From Murray and Correia (2010).

Let i denote the *inclination*, the (smallest) angle between the orbital plane and the plane of the sky. The line formed by the intersection of the two planes is called the *line of nodes*. The *ascending node* is the point in both planes where the orbit crosses the plane of the sky moving from below to above the plane. The *longitude of the ascending node*, Ω, is the angle between the X-axis and the radius vector to the ascending node. The angle between this radius vector and the periastron of the orbit is called the *argument of periastron*, ω.

You can guess where is the *descending node*, and connect the two nodes with the *line of nodes* (see Fig. 2.3)

Ω lies in the XY plane,
ω lies in the xy plane.

The inclination is always in the range $0 \leq i \leq 180°$. An orbit is said to be *prograde* if $i < 90°$ and *retrograde* if $i \geq 90°$.

We can also write the previously defined longitude of periastron

$$\varpi = \Omega + \omega \tag{2.37}$$

but note that, in general (when $i \neq 0$), the angles Ω and ω lie in different planes so that ϖ is a 'dog-leg' angle.

The coordinates in the (x, y, z) system can be expressed in the (X, Y, Z) system after three rotations:

- a rotation about the z-axis through an angle ω
- a rotation about the x-axis through an angle i
- a rotation about the z-axis through an angle Ω.

We can represent these transformations by 3×3 rotation matrices by an angle ϕ, which we denote by $\mathbf{P}_x(\phi)$ (for the rotation about the x-axis) and $\mathbf{P}_z(\phi)$ (for the rotation about the z-axis):

$$\mathbf{P}_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \qquad \mathbf{P}_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then we have

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{P}_z(\Omega)\, \mathbf{P}_x(i)\, \mathbf{P}_z(\omega) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{2.38}$$

Using the expressions for the position of the planet in (x, y, z) coordinates, $x = r\cos f$, $y = r\sin f$ and $z = 0$, we can write the position of the planet in the (X, Y, Z) system:

$$X = r\left(\cos\Omega\cos(\omega + f) - \sin\Omega\sin(\omega + f)\cos I\right) \tag{2.39}$$

$$Y = r\left(\sin\Omega\cos(\omega + f) + \cos\Omega\sin(\omega + f)\cos I\right) \tag{2.40}$$

$$Z = r\sin(\omega + f)\sin I\,. \tag{2.41}$$

If we were interested in deriving the radial velocity *of the planet*, we could do it by calculating $\dot{Z}$ from Eq. (2.41). But we want to arrive at the expression for the radial velocity *of the star*. To do that, we will have to calculate first the motion of the center of mass.

The position vector of the center of mass of the system is

$$\mathbf{R} = \frac{m_1\mathbf{r_1} + m_2\mathbf{r_2}}{m_1 + m_2}\,. \tag{2.42}$$

From Eqs. (2.2) and (2.3) we have

Eqs. (2.2) and (2.3) tell us that $m_1\,\ddot{\mathbf{r}}_1 = -m_2\,\ddot{\mathbf{r}}_2$

$$\ddot{\mathbf{R}} = \frac{m_1\ddot{\mathbf{r}}_1 + m_2\ddot{\mathbf{r}}_2}{m_1 + m_2} = 0\,, \tag{2.43}$$

and by direct integration $\dot{\mathbf{R}} = \mathbf{V} =$ constant. These equations imply that either (i) the center of mass is stationary ($\mathbf{V} = 0$), or (ii) it is moving with a constant velocity ($\mathbf{V} \neq 0$) in a straight line with respect

to the origin O.

If we write $\mathbf{R}_1 = \mathbf{r}_1 - \mathbf{R}$ and $\mathbf{R}_2 = \mathbf{r}_2 - \mathbf{R}$, then

$$m_1\mathbf{R}_1 + m_2\mathbf{R}_2 = 0 \tag{2.44}$$

which implies that $\mathbf{R}_1$ is always in the opposite direction to $\mathbf{R}_2$, and hence that the center of mass is always on the line joining m_1 and m_2. Therefore

$$R_1 + R_2 = r\,, \tag{2.45}$$

where r is the separation of m_1 and m_2, and the distances of the star and planet from their common center of mass are related by $m_1R_1 = -m_2R_2$ (from Eq. (2.44)). Hence

$$R_1 = \frac{m_2}{m_1 + m_2}\,r \quad \text{and} \quad R_2 = -\frac{m_1}{m_1 + m_2}\,r\,. \tag{2.46}$$

Therefore both star and planet orbit the centre of mass of the system in an ellipse with the same eccentricity but the semi-major axes is reduced in scale by a factor

$$a_1 = \frac{m_2}{m_1 + m_2}\,a \quad \text{and} \quad a_2 = \frac{m_1}{m_1 + m_2}\,a\,. \tag{2.47}$$

The orbital periods of the two must each be equal to P and therefore the two mean motions are also equal ($n_1 = n_2 = n$), although the semi-major axes are not. The periapses of the two orbits differ by π.

2.1.3 *The radial velocity equation*

We are nearly there, with all the equations ready to derive the expression for the radial velocity of the star, v_r.

We note again that the plane (X, Y) is taken as the plane of the sky, perpendicular to the line of sight, and the Z-axis is oriented towards the observer. Thus, the radial velocity of the star is simply given by the projection of the velocity vector on the line of sight. Since $\mathbf{r_1} = \mathbf{R} + \mathbf{R_1}$ this gives

$$v_r = \dot{\mathbf{r}}_1 \cdot \hat{\mathbf{Z}} = V_Z + \frac{m_2}{m_1 + m_2}\dot{Z}\,, \tag{2.48}$$

where $V_Z = \mathbf{V} \cdot \hat{\mathbf{Z}}$ is the proper motion of the barycenter and $\dot{Z}$ can be obtained directly from Eq. (2.41):

$$\dot{Z} = \dot{r}\sin(\omega + f)\sin i + r\dot{f}\cos(\omega + f)\sin i\,, \tag{2.49}$$

or, making use of Eqs. (2.25) and (2.26),

$$\dot{Z} = \frac{n\,a\,\sin i}{\sqrt{1 - e^2}}\left[\cos(\omega + f) + e\cos\omega\right]\,. \tag{2.50}$$

We can at last write

$$v_r = V_Z + K\left[\cos(\omega + f) + e\cos\omega\right] , \tag{2.51}$$

where, after using Eq. (2.47),

$$K = \frac{2\pi}{P}\frac{a_1 \sin i}{\sqrt{1-e^2}} . \tag{2.52}$$

We shall refer to Eq. (2.51), when seen as a function of time, as a Keplerian (function). Using Kepler's third law, Eq. (2.21), written as a function of the star's semi-major axis, we can get a more explicit expression for K:

$$K = \left(\frac{2\pi\mathcal{G}}{P}\right)^{1/3} \frac{m_2 \sin i}{(m_1+m_2)^{2/3}} \frac{1}{\sqrt{1-e^2}}. \tag{2.53}$$

A planet with $m_2 \ll m_1$ in a circular orbit ($e = 0$), imprints sinusoidal radial velocity variations on its host star, with a semi-amplitude

The condition $m_2 \ll m_1$ is nearly always a very good approximation. For example, $M_{\rm Jup} \sim 1/1000\ M_\odot$.

$$\left(\frac{K}{\mathrm{m{\cdot}s^{-1}}}\right) = 28.4 \left(\frac{P}{1\,\mathrm{yr}}\right)^{-1/3} \left(\frac{m_2 \sin i}{M_{\rm Jup}}\right) \left(\frac{m_1}{M_\odot}\right)^{-2/3} \tag{2.54}$$

For Jupiter, which has an orbital period of 11.9 years, this amounts to $12.5\,\mathrm{m{\cdot}s^{-1}}$, and for the Earth a mere $0.09\,\mathrm{m{\cdot}s^{-1}}$.

The fastest human on record is able to run at $12.4\,\mathrm{m{\cdot}s^{-1}}$ and the Galápagos giant tortoise moves at about $0.08\,\mathrm{m{\cdot}s^{-1}}$.

* ◯ *

As it may already be clear from the previous equations and from Fig. 2.3, RV measurements alone cannot determine the angle Ω, and the mass of the planet and inclination i are intrinsically degenerate. As a consequence, a RV dataset can provide only five observables for each orbiting planet being considered: P, e, t_0, ω, and K. Given a set of values for these observables, and assuming $m_2 \ll m_1$, we can get the mass of the planet, up to the $\sin i$ factor (Torres et al., 2008):

Even if the parametrization can change (see e.g. Ford, 2005; Beaugé et al., 2012), for example by replacing t_0 with the mean anomaly M, only five independent observables are available.

$$\left(\frac{m_2 \sin i}{M_{\rm Jup}}\right) = 4.919\times10^{-3} \left(\frac{K}{\mathrm{m{\cdot}s^{-1}}}\right) \left(1-e^2\right)^{1/2} \left(\frac{P}{\mathrm{days}}\right)^{1/3} \left(\frac{m_1}{M_\odot}\right)^{2/3} \tag{2.55}$$

and the semi-major axis of the planet's orbit (e.g. Baluev, 2013):

Note the missing 1/3 exponent in (Baluev, 2013), their Eq. (5).

$$\left(\frac{a_2}{\mathrm{AU}}\right) = 1.957\times10^{-2} \left(\frac{P}{\mathrm{days}}\right)^{2/3} \left(\frac{m_1}{M_\odot}\right)^{1/3} . \tag{2.56}$$

By itself, a periodic RV modulation is only sufficient to infer the presence of a companion, but it is not sufficient to infer its true mass: small values of K can indicate either a low-mass planet on an orbit with an inclination close to 90 degrees, or an object of higher mass with a smaller orbital inclination.

Throughout the Thesis, when I refer to a planet mass determined

from RV observations, this should be understood as a minimum mass $m_p \sin i$ and not a true mass m_p, even if not clearly stated. Nevertheless, the impact of the $\sin i$ factor is usually considered to be mild. If we assume a random orientation of the orbital plane, the distribution of orbital inclinations is proportional to $\sin i$ (e.g. Ho and Turner, 2011). Therefore, the probability of having an inclination in the range $[0, i_x]$ is

$$\int_0^{i_x} \sin i \, \mathrm{d}i = 1 - \cos i_x$$

which means that only 13.4% of all systems will have $i < 30^\circ$, corresponding to $\sin i < 0.5$ and therefore a true mass more than twice the measured minimum mass. See however Ho and Turner (2011) and Lopez and Jenkins (2012) for a more in-depth discussion.

2.1.4 *The effect of multiple planets*

The first assumption we made, right at the start of section 2.1.1, was that there were *two* bodies, a star and a planet. What happens when a star hosts $N_p > 1$ planets? The gravitational interactions between the planets can affect their orbits in quite complex ways, and change the radial velocity of the star.

When more than one planet orbits the star, the equation of motion for planet i (with mass m_i) changes from Eq. (2.4) into

$$\ddot{\mathbf{r}} = -\mathcal{G}\,(m_0 + m_i)\frac{\mathbf{r}_i}{r_i^3} + \mathcal{G}\sum_{j=1; j\neq i}^{N_p} m_j \left(\frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3} - \frac{\mathbf{r}_j}{r_j^3} \right), \tag{2.57}$$

Here we are neglecting the effects of relativity and assuming spherical bodies. See, e.g., Fabrycky (2010).

where each of the coordinates $\mathbf{r}_i$ is referred to the central star, of mass m_0. The sum over each of the other planets contains a term due to the direct gravitational force and one due to the forces that cause the star to accelerate.

Unlike before, there is no analytical solution to Eq. (2.57), and the calculation of the orbits has to be done by numerical integration. Another option when considering multiple planets is to ignore the mutual gravitational interactions and approximate the total radial velocity signal as a linear sum of Keplerian functions from each planet

$$v_r = \sum_i v_{r,i} \tag{2.58}$$

where i indexes the planets.

Deviations from this simple sum-of-Keplerians model are often small when compared with the amplitude of the Keplerians and the observational uncertainties. Nevertheless, in some exoplanet systems which are in or near a mean motion resonance, a so-called *self-consistent* N-body model is required to accurately describe the observations (see for example Correia et al., 2010; Rivera et al., 2010;

Cochran et al., 2011; Johnson et al., 2011; Fabrycky et al., 2012; Nelson et al., 2014).

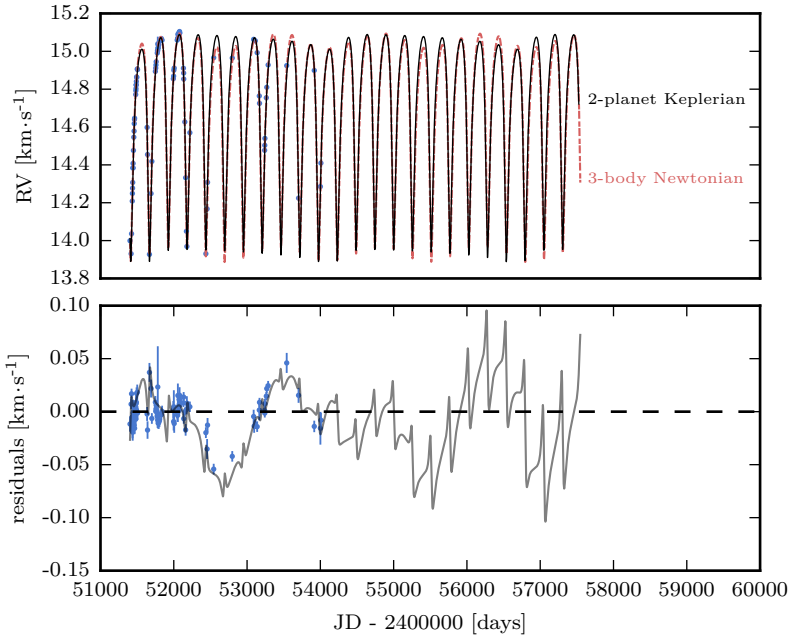


Figure 2.4: Analysis of 92 RV observations of HD 202206 from the ELODIE spectrograh. The data was presented in Couetdic et al. (2010). The top panel shows the RV observations together with a 2-planet Keplerian fit (which neglects the gravitational interactions between the planets) and a 3-body Newtonian fit (which takes into account the interactions). The bottom panel shows the residuals from the Keplerian fit, which follow closely the difference between the sum-of-Keplerians signal and the Newtonian signal.

As a a brief example, I show in Fig. 2.4 the analysis of ELODIE data for HD 202206, where a massive inner planet induces large perturbations in the orbit of an outer one (Correia et al., 2005; Goździewski et al., 2006; Couetdic et al., 2010). I consider a 2-planet Keplerian fit to the data, using the sum-of-Keplerians model, and compare it to a 3-body (2 planets and the star) Newtonian model. The residuals from the Keplerian fit are well aligned with the difference between the sum-of-Keplerians signal and the Newtonian signal. In this case, the perturbations are detectable even at the level of precision of ELODIE (the average errorbar is $\sim 8\,\mathrm{m\cdot s^{-1}}$).

When planet-planet interactions are present and detectable, they can provide a very detailed view of the three-dimensional orbits. In some cases, the orbital plane inclinations can be constrained and the true masses of the planets determined. These systems also present excellent benchmarks for theories of planet formation and migration, as was alluded to in Chapter 1.

2.2 MEASURING RADIAL VELOCITIES

Now we know that a star wobbles around the planetary system center of mass, and the velocity – along our line of sight – with which it moves. How can we *measure* this radial velocity? Any astronomer will answer immediately: measure the Doppler shift of the stellar spectral lines.[2] But the devil is in the details.

[2] Maybe Dravins et al. (1999) would give a longer answer.

2.2.1 *The Doppler effect*

We are back in 1904, knowing nothing about relativity (Einstein, 1905). A photon with wavelength λ_0 in the rest frame of the source (the star) will be detected at a different wavelength λ in the rest frame of an observer (our spectrograph) which is moving with respect to the source (or vice-versa, if you wish). The difference in wavelengths, $\Delta\lambda$, will be given by (Doppler, 1842)

$$\frac{\Delta\lambda}{\lambda_0} = \frac{v}{c}\,, \tag{2.59}$$

where v is the velocity with which the source is moving relative to the observer and c is the speed of light.

The myriad effects caused by special and general relativity is fascinating. Look at Lindegren and Dravins (2003), Lovis and Seager (2010), Cegla et al. (2012), and Wright and Eastman (2014).

The wavelength difference caused by the Doppler effect can indeed be measured on the many spectral lines present in the spectrum of the star (Fig. 2.5). But achieving a high precision when measuring these Doppler shifts is still a challenging task.

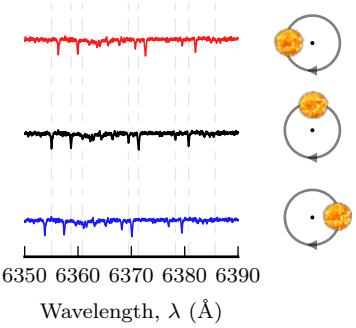


Figure 2.5: Illustration of the Doppler effect on a stellar spectrum. As the star orbits the centre of mass, the spectrum is red-shifted and blue-shifted relative to the position of reference spectral lines. The observer would be located here.

A value $v = 1\,\mathrm{m\cdot s^{-1}}$ in Eq. (2.59) corresponds to wavelength shifts of about 10^{-5}Å in the visible ($\lambda_0 \sim 500$nm). For an instrument with a spectral resolution $R = 100\,000$, this represents 1/3000 of the width of a typical line, or about 1/1000 of a CCD pixel in the detector (here we used the expression $R = \lambda/\Delta\lambda = c/v$, considering $\Delta\lambda$ to be the resolution element or the width of a line, and that a typical high-resolution spectrograph projects the spectra on a detector reaching a sampling of 3 pixels per resolution element e.g. Mayor et al., 2003).

In order to achieve a high-level RV precision, one must control two main sources of noise: statistical errors coming from the RV measurement itself and systematic errors induced by the spectrograph. This requires a stabilized spectrograph, precisely calibrated in wavelength, and the combination of thousands of spectral lines (Pepe et al., 2014a, and references therein).

In any spectrograph, the contribution of instrumental photon noise to the error in the RV measurement can be expressed as (Hatzes and Cochran 1992; see also Bouchy et al. 2001 for a different formulation):

$$\sigma_{\mathrm{RV}} \propto \frac{1}{\sqrt{F}\cdot\sqrt{\Lambda}\cdot R^{1.5}}, \tag{2.60}$$

where F is the flux and Λ is the wavelength range. From this formula,

one sees that the instrumental requirements for high-precision radial velocities are very high spectral resolution ($R > 50\,000$), a large spectral coverage and high efficiency. Only cross-dispersed echelle spectrographs are capable of delivering all these requirements (while also keeping the instrument within a reasonable size).

2.2.2 *Spectra of cool stars*

Because stars emit most of their light between the ultraviolet and mid-infrared domains, and because those spectral regions themselves are not accessible from the ground, radial velocity measurements are usually made using spectral lines within the visible and near-infrared.

Across the HR diagram, stellar spectra change quite considerably (see Fig. 2.6 and, e.g., Gray and Corbally 2009). Hot stars (with $T_{\rm eff} > 10\,000K$) show almost no spectral absorption lines in the visible and near-IR. At lower temperatures ($T_{\rm eff} < 3500K$) the lines are densely packed, less contrasted and overlapping because of molecular bands. These stars are also intrinsically faint and emit most of their light in the infrared (e.g. Gray, 2005).

Solar-type stars and M dwarfs, those with spectral types from $F5$ to $M5$, are therefore the best targets for precise radial velocity determinations. In the main sequence, this corresponds to masses between $0.1M_\odot$ and $1.5M_\odot$. These stars have thousands of absorption lines in their spectra, which are produced by many different chemical elements, with Iron lines being the most common.

To increase the precision with which we measure radial velocities, we need to use as many spectral lines as possible. However, strongly saturated lines, such as the Hydrogen $H\alpha$, the Calcium H&K lines and the Sodium D doublet, should be avoided because of their broad wings and variable chromospheric emission in their cores (which is in fact used to trace stellar activity, see Section 3.2).

We briefly saw before how the RV precision depends on characteristics of the spectrograph (Eq. 2.60), and we can similarly relate it to the shapes of spectral lines. Assuming approximately Gaussian shapes for the lines, it can be shown (e.g. Bouchy et al., 2001) that

$$\sigma_{\rm RV} \sim \frac{\sqrt{\rm FWHM}}{C \cdot \rm SNR}, \tag{2.61}$$

where FWHM is the full width at half maximum of the line, C is its contrast (the depth divided by the continuum level), and SNR is the signal-to-noise ratio in the continuum. Eq. (2.61) can be understood intuitively, as the deeper and narrower the line, the better defined its centroid will be, allowing for a more precise measurement of the radial velocity.

Another very important aspect determining the RV precision is

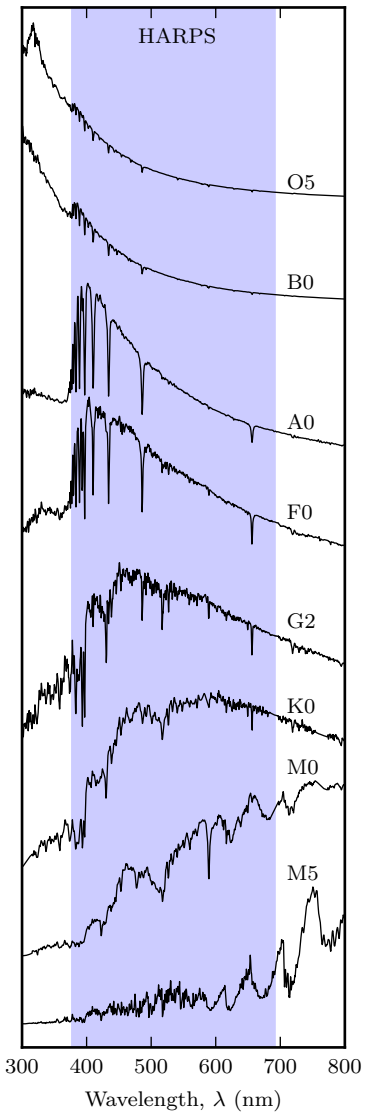


Figure 2.6: Spectra of stars belonging to different spectral types. The blue region is the wavelength coverage of the HARPS spectrograph (378nm - 691nm), see Section 2.2.5.

controlled by the rotation of the star. To see this, consider an approximation to the area of a Gaussian spectral line, given by $C \times \mathrm{FWHM}$. The lines of a rotating star are broadened by a rotation kernel of unit area (Gray, 2005). Therefore this rotation kernel changes the FWHM and C of a spectral line, but the value $C \times \mathrm{FWHM}$ remains the same. From Eq. (2.61) we can then obtain that σ_{RV} will depend on the rotation velocity of the star, $v \sin i$, to the power 1.5, the same dependence as with the spectral resolution.

2.2.3 *The cross-correlation function*

Given a wavelength-calibrated spectrum with thousands of lines, one could fit Gaussian profiles to each line, measure how much their centres deviate from the rest-frame wavelengths, plug those values into Eq. (2.59) and calculate their corresponding radial velocities. A weighted average of these values would give us the radial velocity of the star.

That sounds OK, but slightly inefficient, both in terms of computation and also on the calculation of the uncertainty in the measured RV, which would have to be propagated from each individual line. Let us instead try to concentrate all the information contained in the spectrum on a single line and measure *its* radial velocity.

We do this with a variation of the cross-correlation method. In general, cross-correlation is a way of computing the global shift between two similar signals. In our case, one signal is the spectrum, the other is a binary transmission function with ones at the (rest-frame) position of stellar lines and zeros elsewhere. The global shift will be the radial velocity.

This procedure is nearly optimal in extracting all the Doppler information from the spectrum and concentrating it into a single *cross-correlation function* (CCF), an average line (e.g. Bouchy et al., 2001). It is not optimal because the RV information of a line is proportional to its depth (see Eq. 2.61). But this is easy to fix: the binary mask can contain the line depths instead of ones. Then, the cross-correlation of the spectrum $S[\lambda_i]$ with the weighted mask $M[\lambda_i]$ gives us the optimal cross-correlation function (Pepe et al., 2002a):

$$\mathrm{CCF}(v) = \sum_i S[\lambda_i] \cdot M[\lambda_i(1 + v/c)] \tag{2.62}$$

where $M[\lambda_i(1 + v/c)]$ is the mask Doppler-shifted by an RV value v. Note that, for a given star, all lines used in the mask can be assumed to have the same FWHM (as defined by the spectrograph and the rotation of the star in the case of a non-saturated line), and so line depth is the only parameter to include in the weights.

Finally, we can measure the radial velocity by fitting a suitable

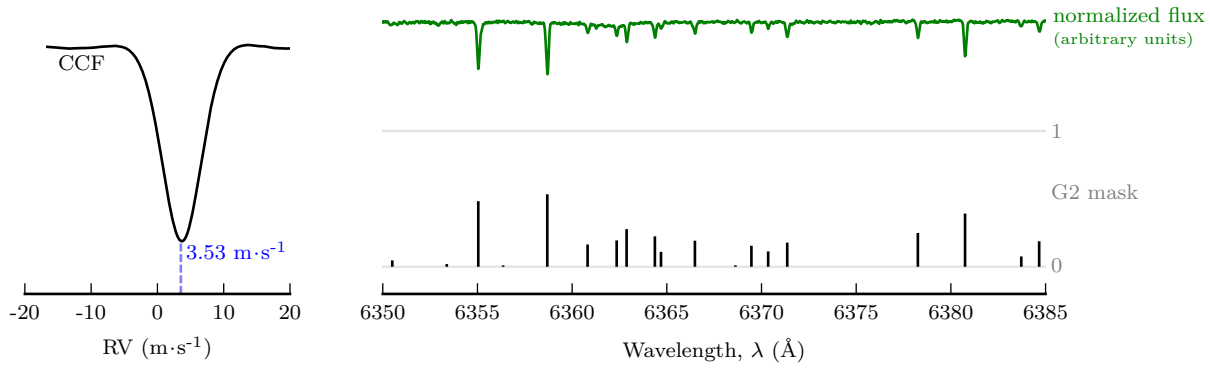


Figure 2.7: Cross-correlating the observed spectrum of a star (green) with a mask produces the cross-correlation function, on the left. This average spectral line is offset from 0 by the radial velocity of the star. Note that the figure only shows a (very small) portion of the stellar spectrum, the full mask includes around 4000 lines.

model to the CCF, usually a Gaussian function for slowly-rotating stars (see Fig. 2.7). In typical observed CCFs, deviations from the Gaussian shape are much smaller than 1%. However, they are asymmetric. We will come back to this important point in Chapter 3, when we discuss the effects of stellar activity.

We could in principle create a mask for each star, against which we would correlate any observed spectrum. But that would imply a large observational effort, and usually a single mask for each main spectral type (F, G, K) is sufficient.

For later spectral types, i.e. M dwarfs, the spectra become overpopulated with lines, such that the average distance between lines is smaller than the resolution of the spectrograph, creating line blends (see e.g. Figueira et al., 2016). In these situations, the CCF does not deliver the best precision, and it is often preferable to cross-correlate the spectrum with a template, either derived from a theoretical model or from an average spectrum calculated directly from the observations (e.g. Anglada-Escudé and Butler, 2012; Astudillo-Defru et al., 2015).

2.2.4 *The impact of stellar metallicity*

Spectral absorption lines are the most important source of information to establish the chemical composition of the stellar atmosphere. Line strengths are heavily dependent on temperature, which dictates the relative populations of the energy levels, and to a lesser extent on pressure. But for any given values of temperature and pressure, the strength of an unsaturated absorption line is nevertheless expected to increase with the number of atoms of a given element (Gray, 2005). This means that lower metallicity stars will show weaker absorption lines, i.e., with smaller depths.

Stellar metallicity is usually represented by the Iron abundance, denoted [Fe/H]. Iron is not the most abundant heavy element, but its abundance is easy to measure because of many absorption lines in the optical spectrum. It is also usually assumed that the abundance of

What does [Fe/H] mean? The definition of this scale is

$$[X/H] = \log_{10}\left(\frac{n_X}{n_H}\right)_* - \log_{10}\left(\frac{n_X}{n_H}\right)_\odot$$

where n_X is the number of atoms from element X and thus n_H is the number of Hydrogen atoms.

Iron scales with the abundance of other metals and hence with overall metallicity (however, metal-poor stars have a rather different relative abundance pattern of elements, see e.g. Adibekyan et al. 2012b).

Fig. 2.8 shows a small part of the spectra of two stars with similar effective temperatures, one with [Fe/H] = −1.39 (only 4% of the amount of Iron in the Sun) and the other with [Fe/H] = −0.05 (about 90% of the solar metal abundance). Note the much weaker absorption lines in the lower metallicity star.

Since the precise determination of radial velocities relies on combining the information from many spectral lines, and this information depends on the line depth, it is clear that the RV precision obtained for metal-poor stars will be inherently lower. We can sometimes compensate for this effect by taking longer exposures on more metal-poor stars, therefore obtaining spectra with higher SNR.

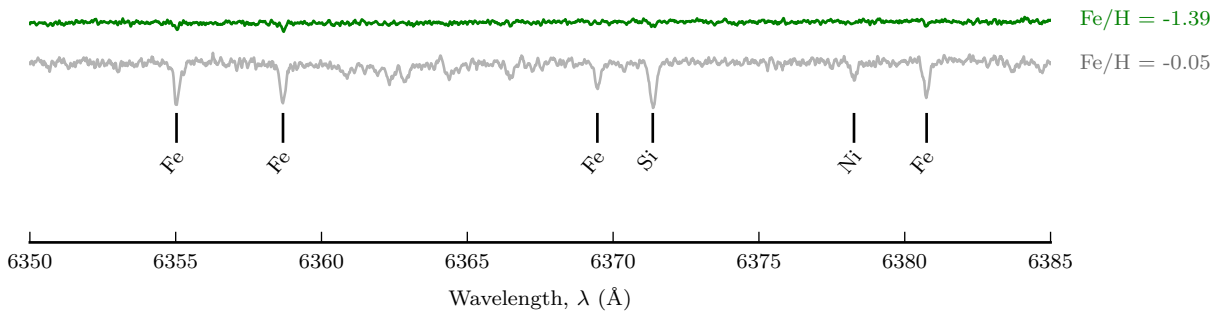


Figure 2.8: Spectra of two stars with different metallicities.

2.2.5 *HARPS, the planet hunter*

This brief section will introduce the HARPS instrument, used to collect most of the RV data analysed in this Thesis. The High Accuracy Radial velocity Planet Searcher (Pepe et al., 2002b; Mayor et al., 2003) is currently one of the most precise spectrographs dedicated to measuring radial velocities. It is a stabilized cross-dispersed echelle spectrograph, fiber-fed by the Cassegrain focus of the ESO 3.6m telescope, at the La Silla observatory in Chile. It achieves a resolution of 115 000 and the spectral coverage is from 378 nm to 691 nm.

Two optical fibers go into HARPS: one collects the stellar light, and the other is used to either record simultaneously a Th-Ar reference spectrum or the background sky. This is the so-called simultaneous-reference technique (e.g. Baranne et al., 1996; Fischer et al., 2016). The precision achieved by HARPS is due, in part, to this simultaneous calibration. The reference spectrum coming from the second fiber helps in determining a wavelength calibration, and how this calibration changes with time. This allows a precise measurement of any RV

drift that might occur during a night of observations.

Any such RV drift is nevertheless expected to be small because HARPS operates under vacuum, and is pressure and temperature controlled at 0.01 mBar and 0.01 K, respectively (Mayor et al., 2003). This leads to controlled variations of the instrumental profile.

HARPS is capable of a long-term RV precision of $1\,\mathrm{m\cdot s^{-1}}$ (Rupprecht et al., 2004; Lovis et al., 2006). In its 13 years of operations, it has single–handedly discovered hundreds of exoplanets (see for example Mayor et al., 2011), including some with masses close to that of the Earth. As part of a recent upgrade, in June 2015, a new set of octagonal optical fibers has been installed (Lo Curto et al., 2015). These fibers substantially improve the throughput of the instrument and decrease its sensitivity to de-centering and de-focusing. However, an RV offset of the order of $15\,\mathrm{m\cdot s^{-1}}$ was introduced with this change. In section 4.3.1, we will show how to deal with this offset.

2.3 ANALYSIS OF RADIAL-VELOCITY DATASETS

In the previous Sections we saw how spectral absorption lines can be used to measure the radial velocity of a star. We now have to analyse a time series of RVs in our quest for exoplanets.

But before we start with the analysis, we should describe what a typical dataset of RV measurements looks like. This will be a high-level description, basically detailing the output of the HARPS Data Reduction Software (DRS), the pipeline used to reduce HARPS data.

Each observation, indexed by k, is identified with a timestamp t_k, which describes when the observation occurred. This is usually presented in the Modified Julian Date (MJD), which is obtained by subtracting 2 400 000.5 days from the Julian date (see Montenbruck and Pfleger, 2000). There are a couple of practical considerations to take into account here. First, the effective time of a lengthy observation is determined using an exposure meter and is set to the time of arrival of the median photon (e.g. Wright and Eastman, 2014). Also, when the observations are binned at individual nights (see Section 3.2.1) the corresponding timestamp is the weighted average of the original times, the weight being the inverse quadratic sum of the variances associated to each spectra.

HARPS started observing on October 1, 2003, that is MJD 52913. This chapter is being written on June 8, 2017, that is MJD 57912.

At each timestamp, the dataset contains the value of the barycentric stellar radial-velocity, $v_{\mathrm{obs}}(t_k)$, and associated uncertainty, σ_k. These are usually measured in $\mathrm{km\cdot s^{-1}}$, although in some cases throughout the Thesis you will see them presented in $\mathrm{m\cdot s^{-1}}$. The RV uncertainty contains the contributions from the photon noise, the instrumental drift, and the wavelength calibration.

file: `starid_harps.rdb`

`time`	`vrad`	`svrad`	← in the file
t_k	$v_{\mathrm{obs}}(t_k)$	σ_k	← in the Thesis
⋮	⋮	⋮	
⋮	⋮	⋮	k ↓
⋮	⋮	⋮	

The HARPS pipeline also delivers a number of other quantities,

the principal ones being the line-profile indicators obtained from the CCF. These will be discussed in Chapter 3. For now, we continue with the analysis of the RV time series.

2.3.1 *A model for planetary companions*

Let us recall the expression for the radial velocity variations of a star with an orbiting planet, Eq. (2.51):

$$v_r(t) = v_{\rm sys} + K\{\cos\left[\omega + f(t)\right] + e\cos\omega\}\,, \tag{2.63}$$

where we have changed the symbol V_Z, representing the systemic velocity of the star along the line of sight, to $v_{\rm sys}$. For a star being perturbed by multiple planets, and ignoring planet-planet interactions, we have:

$$v_r(t) = v_{\rm sys} + \sum_i K_i\{\cos\left[\omega_i + f_i(t)\right] + e_i\cos\omega_i\}\,, \tag{2.64}$$

where i indexes the planets.

These equations represent our model[3] for planetary RV signals, and we will use $v_{\rm model}$ to name them. Because the model depends on the set of parameters $\boldsymbol{\theta}$ and on time, we write $v_{\rm model}(\boldsymbol{\theta}, t)$. For clarity, we can list the full set of parameters included in $\boldsymbol{\theta}$, containing the orbital parameters and the systemic velocity:

$$\boldsymbol{\theta} = \{P_i, K_i, e_i, \omega_i, \phi_i\}\,,\, v_{\rm sys}$$

With a model at hand, we can proceed to *fit* a RV dataset.

[3] Or at least its deterministic part.

Note that we chose to parametrize the Keplerian in terms of a phase ϕ, instead of the time of periastron T_0. They are related by

$$T_0 = t_1 - P\,\frac{\phi}{2\,\pi}$$

where t_1 is the time of first observation.

2.3.2 *Maximum-likelihood approach*

Since the uncertainties from individual RV observations are expected to follow a normal distribution, we are prompted to evaluate the goodness-of-fit for a given model with the well-known χ^2 statistic:

$$\chi^2 = \sum_k \frac{\left[v_{\rm model}(\boldsymbol{\theta}, t_k) - v_{\rm obs}(t_k)\right]^2}{\sigma_k^2} \tag{2.65}$$

The maximum likelihood estimate (MLE) for the set of model parameters $\boldsymbol{\theta}$ is obtained by finding the minimum χ^2. That is, we find the value of $\boldsymbol{\theta}$ that minimizes Eq. (2.65). These are the parameters typically reported for the "best-fit" model.

Methodologically, finding the MLE requires a minimization algorithm. The Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) seems to be the most popular in the exoplanet literature. From an initial guess for the parameters, it quickly converges to a local minimum of the χ^2.

To come up with a reasonable initial guess for the orbital period

(which turns out to be the most difficult parameter), one usually calculates the Lomb-Scargle periodogram of the RV observations (Lomb, 1976; Scargle, 1982), or one of its generalizations, the so-called Generalised Lomb-Scargle (GLS; Zechmeister and Kürster, 2009) or its Bayesian version, the BGLS (Mortier et al., 2015), for example. The periodogram is certainly an interesting subject on itself, but exploring it is outside the scope of this Thesis. See VanderPlas (2017) for a very clear discussion.

It is unfortunate that the $\chi^2(\boldsymbol{\theta})$ surface can sometimes present a large number of local minima, specially for multiple-planet models. This can make the minimization algorithm particularly sensitive to the initial guesses and hinder convergence.

To remedy this issue, a number of *global* minimization algorithms have been proposed to reduce the risk of getting stuck in a local minimum and the dependence on the initial guesses. In the literature, Simulated Annealing and Genetic Algorithms are the most common (e.g. Stepinski et al., 2000; Iglesias-Marzoa et al., 2015).

Once a (hopefully global) minimum is found in the $\chi^2(\boldsymbol{\theta})$ surface, we are left with the task of determining the uncertainties in the estimated orbital parameters. Again, the classical approach branches into (at least) two main methods: constant χ^2 boundaries and resampling techniques: (*i*) given the minimum value $\chi^2_{\rm min}$, we can vary the model parameters slightly around their optimal values such that χ^2 changes by less than 1 or (*ii*) we can simulate repeated measurements of the data (assuming that we know the error distribution) and perform the χ^2 minimisation for each realisation. The distribution of the best-fit parameters from all those fits to mock data gives us an estimate for the uncertainties in the parameters.

In Fig. 2.9, I show the results of the analysis of HARPS data for HD 72659. The dataset contains 46 RV observations of this star. Using the Levenberg-Marquardt algorithm,[4] I found the minimum χ^2 value, and the corresponding best-fit parameters. The numerical value was found to be $\chi^2 = 197.14$.

Varying the orbital period in a small range around the fitted value, the minimum χ^2 values were found, by optimizing for the remaining parameters,[5] so that the boundary where the χ^2 changes by 1 could be determined. This boundary corresponds to a first estimate of the 1σ uncertainty in the orbital period. As shown in Fig. 2.9, top right panel, this estimate is $P = 3016.68 \pm 18.23$ days.

To illustrate the data resampling technique, I generated a series of normally-distributed random numbers, with variances given by the uncertainty of each data point, σ_k^2. These random numbers were added to the observed RVs[6] and a new fit was performed (by minimising again the χ^2). Note that all parameters are free in each of

[4] With initial conditions close to the values found by Wittenmyer et al. (2007).

[5] The orbital period P is kept fixed in a given range around the best-fit value, and the other parameters are optimized for each value of P.

[6] The observational uncertainties were not altered.

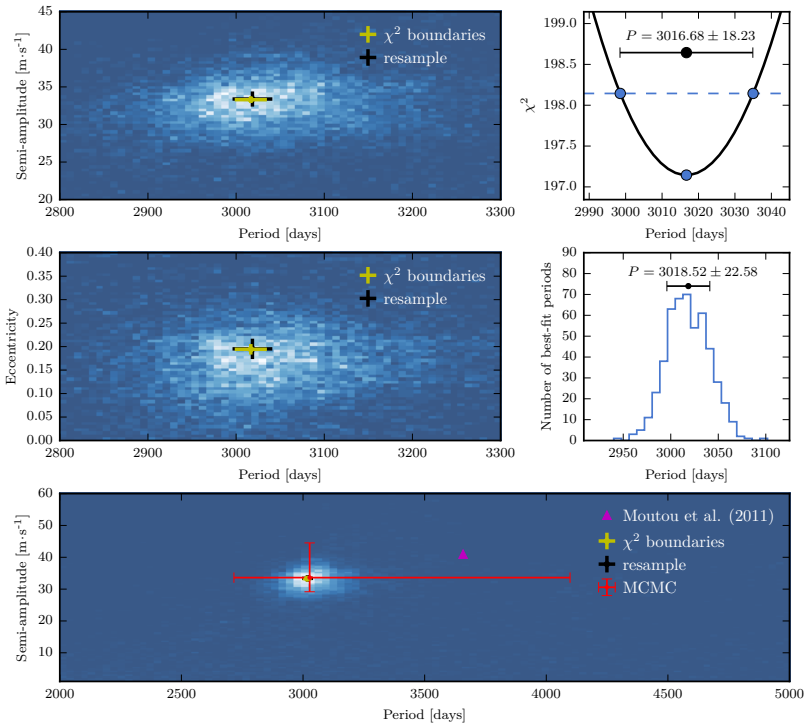


Figure 2.9: Analysis of HARPS data for HD 72659. The top and middle right panels show a visual demonstration of the χ^2 boundaries and resampling techniques for error estimation. On the left, the uncertainties are shown in the K-vs-P and e-vs-P spaces, superposed on the posterior distribution obtained from an MCMC analysis. The bottom panel shows a zoom-out of the K-vs-P distribution, with the full extent of the MCMC error bars and the value obtained by Moutou et al. (2011) from combined HARPS and Keck/HIRES data.

the new fits. The distribution of orbital periods found for all the fits is shown in the middle right panel of Fig. 2.9. The mean and standard deviation of this distribution provide an estimate for the orbital period: $P = 3018.52 \pm 22.58$ days.

The bottom and middle left panels of Fig. 2.9 show the comparison of these two estimates with the results from an MCMC run on the same dataset (more details on the model will be presented in the following Sections). The 2-dimensional distribution shows an histogram of samples from the posterior distribution for the semi-amplitude and the eccentricity versus the orbital period. Yellow and black errorbars represent the estimates for these three parameters obtained using the χ^2 boundaries and the resample methods, respectively. The uncertainty estimates for the semi-amplitude and eccentricity were obtained exactly in the same manner as for the orbital period.

Finally, the bottom panel in Fig. 2.9 is a zoom out of the top left panel, showing the complete posterior distribution for the semi-amplitude and orbital period on a wider range. At this scale, the error estimates from χ^2 boundaries and resampling are nearly invisible. The red errorbar is the error estimate from the MCMC, calculated as the median of the posterior distribution for 1-planet solutions and the 68% quantiles of the complete posterior distribution. The magenta triangle is the orbital parameters derived by Moutou et al. (2011), who used both HARPS and Keck data in a combined analysis.

If (and only if) we consider the estimates derived from the com-

bined dataset to be more reliable and correct, it becomes clear that only the MCMC error estimates come close to truly express our uncertainty. The classical error estimates are severely underestimated. This happens for several reasons, the main one being that the classic uncertainties for each parameter mostly ignore the variation of all the other parameters. It reflects the difference between an estimate of the distribution *conditional* on specific values of the other parameters, or the distribution *marginalised* over the other parameters.

* ◯ *

After having obtained the best-fit parameters and their associated uncertainties, we now have to specify *how well* this model performs at fitting the data, when compared with other models. Put another way, given two models with different number of parameters, e.g. a 1-planet model and a 2-planet model, how can we decide which one provides the best fit?

The easy answer would be to choose the model with the lowest value of χ^2. But it is easy to see that the value of χ^2 can be made arbitrarily small just by increasing the number of parameters in the model. We would always end up choosing a model with too many planets because adding one extra Keplerian will always lower the χ^2.

So the *number* of parameters in each model is an important variable, which can be considered in the goodness-of-fit and in comparing different models. The *reduced* χ^2 is often used:

$$\chi^2_r = \frac{\chi^2}{N-p}, \tag{2.66}$$

where p is the number of free parameters in the model. A model A is better at fitting the data than a model B, if $\chi^2_{r,\mathrm{A}}$ is smaller than $\chi^2_{r,\mathrm{B}}$, or if it is closer to 1.

Andrae et al. (2010) discuss some of the pitfalls of using this quantity, the main one being that the factor $N-p$ often does not actually represent what it seeks to represent: the number of *degrees of freedom* of the model. Nevertheless, the reduced χ^2 is widely used for model assessment and comparison.

As mentioned before, the χ^2 and the likelihood are related, in that the best-fit parameters, which minimise the χ^2, maximise the likelihood function. Therefore, to the minimum χ^2 value there is a corresponding numerical value of the likelihood at its maximum point. Let us denote this value by $\hat{L}$. Much alike the reduced χ^2, but with somewhat deeper theoretical grounds, a number of *information criteria* using $\hat{L}$ have been proposed for model comparison. The most widely used are the AIC (Akaike, 1973) and the BIC (Schwarz, 1978):

AIC is almost always expanded to 'Akaike's Information Criterion', even if Akaike (1973) actually proposed to call it simply 'An Information Criterion'. BIC, in turn, stands for 'Bayesian Information Criterion' (Schwarz, 1978).

$$AIC = -2\ln(\hat{L}) + 2\,p \tag{2.67}$$

$$BIC = -2\ln(\hat{L}) + p\ln N. \tag{2.68}$$

In both these quantities, the first terms of the right-hand side are measures of how well the model fits the data, while the second terms can be interpreted as "penalties" for increasing the size of the model (see Burnham and Anderson, 2010).

Note that the penalty terms depend linearly on the number of parameters, meaning that the criticism of Andrae et al. (2010) applies to AIC and BIC in the same way.

In practice, one computes the value of AIC (BIC) for each of the candidate models, selecting that one with the smallest value of AIC (BIC). This model is estimated to be the one "closest" to the unknown reality that generated the data (from the models considered). The scale of the differences between any two AIC (BIC) values provides a measure of significance for the model comparison: AIC (BIC) differences over 10 are usually required (Burnham and Anderson, 2010).

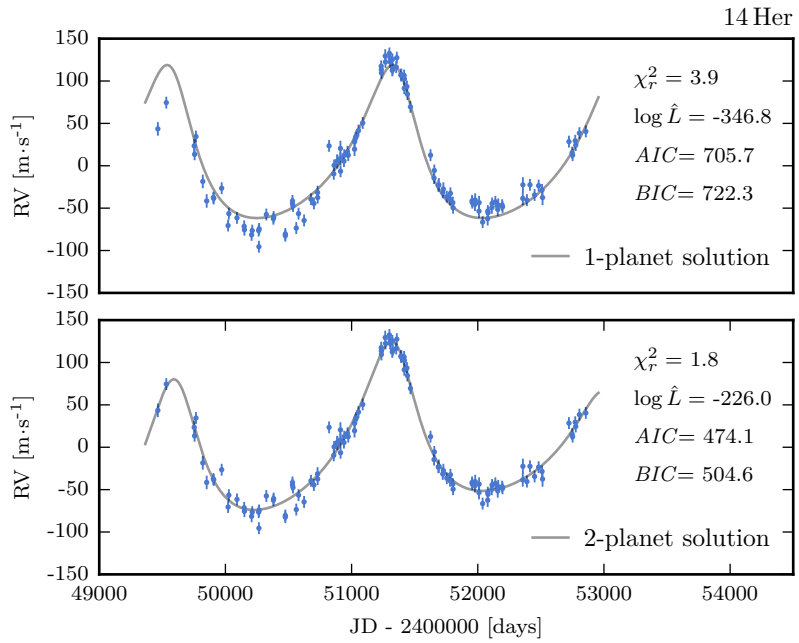


Figure 2.10: ELODIE data and orbital solutions for 14 Her. *Top:* 1-planet solution; *bottom:* 2-planet solution. There are 119 measurements with a mean error bar of 7.2 m·s^{-1}. The fits were calculated by minimizing the χ^2 with the Levenberg-Marquardt algorithm, as implemented in the `scipy` package (Jones et al., 2001). The values of χ^2_r, $\log \hat{L}$, *AIC* and *BIC* are shown for each of the models.

Fig. 2.10 shows the analysis of ELODIE data for 14 Her (HD 145675), a star known to host at least one Jupiter-mass planet. The solutions with one and two planets were found by minimizing the χ^2, and the model comparison diagnostics discussed above, are shown for the two models. The two-planet solution is clearly preferred, by all four quantities (see Wittenmyer et al., 2007).

Now the question is what happens with a 3-planet model? In this case, the χ^2 minimization starts to depend strongly on the initial values of the parameters. Nevertheless, it is possible to find solutions with 3 Keplerian signals with χ^2_r as low as 1.67, AIC = 455.9 and BIC = 500.4. Comparing these values with those in Fig. 2.10 for the 2-planet fit, one could easily argue (if they were using the AIC

for example, since $474.1 - 455.9 > 10$) that the 3-planet model is significantly preferred.

In this case the BIC also prefers the solution with 3 planets but with less significance ($\Delta\text{BIC} = 504.6 - 500.4$).

At this point, presented with the choice between announcing 2 or 3 planets, one could come up with new criteria that would make those 3-planet solutions less credible (if we find them to be less "believable"). For example, we could say that the BIC indeed chooses only 2 planets as the best solution, and we would only announce a new planet when the two information criteria agree. Not only the arbitrariness but the data-dependence of these judgements then becomes clear.

2.3.3 *A critique*

The methods of the previous section are seemingly reliable, relatively easy to understand, and usually quite fast. But they lack the appeal of a complete, self-consistent, coherent theory, which Bayesian probability theory possesses (Jeffreys, 1961; Jaynes, 2003).

There are a few standard objections to classical (maximum likelihood, frequentist, sometimes called orthodox) inference. In his unfinished book, Jaynes (2003, Chapter 16) exposes the arguments that appeal the strongest to me.

The early days of what can be called statistical inference were dominated by, essentially, Bayesian methods, in the works of Laplace, Gauss, Legendre, Poisson, etc. But in the beginning of the 20th century the relationship between statisticians and scientists had evolved into a supplier-client or doctor-patient one. Jaynes argues on what might have been the reason for this: with the lack of unifying principles "the collection of all the empirical procedures a data analyst might need, like the collection of all the logically unrelated medicines [...] that a sick patient might need, was too large for anyone but a dedicated professional to learn" (Jaynes, 2003, p. 492).

The works of Jeffreys (1961) and Jaynes[7] have provided the necessary foundational arguments to Bayesian probability theory, and in turn shown why orthodox statistics works and why it fails.

[7] Together with many other authors, see Savage (1972), Berger (1985), Bernardo and Smith (1994), and Robert et al. (2009, and references therein).

As we saw with the analysis of the RVs of HD 72659, classical methods can very easily underestimate the uncertainties in derived parameters. In addition, a number of different model selection criteria are commonly used, offering little justification. They are also in disagreement more often than not, as in the case of 14 Her. Bayesian model comparison offers one and only one answer to this problem (e.g. Ford and Gregory, 2007).

Because I take the Bayesian choice (Robert 2007; see also Loredo 1994), we will stop here with the discussion of the χ^2, ad-hoc methods of error estimation and of model comparison. Instead, we continue with the definition of a complete probabilistic model for RV observations and its estimation using a custom Monte Carlo algorithm.

2.3.4 *Trans-dimensional inference*

In this section we will build a custom probabilistic framework for the analysis of RV datasets, based on the work of Brewer and Donovan (2015). I will go through the definition of the likelihood, the prior distribution for all the model parameters, and the evidence until, finally, the complete model is presented.

Our problem is the following. There is an unknown number, N_p, of *detectable* planets in the radial velocity dataset of a star. Each of the N_p planets has unknown orbital parameters $\boldsymbol{\theta}$, which include the orbital periods, eccentricities, semi-amplitudes, etc. Given data $\mathcal{D}$, our goal is to infer both N_p and $\{\boldsymbol{\theta}_i\}_{i=1}^{N_p}$.

A number of undetectable planets might nevertheless exist around that star.

It is important to point out that we would like to infer all the parameters *simultaneously*.

Inference, in this setting, means that we use the data $\mathcal{D}$ to improve our knowledge about the parameters N_p and $\boldsymbol{\theta}$, so that one can take a decision related with them. The inference should be able to deal with either N_p or some components of $\boldsymbol{\theta}$ alone: we can ask "What is the value of P_1?" (the orbital period of the first planet) or, for the eccentricities, "Is e_1 larger than e_2?".

Because N_p may be large, it is often useful (though not necessary) to introduce a set of parameters $\boldsymbol{\alpha}$ that will *control* for $\boldsymbol{\theta}$. This basically allows us to input available prior knowledge about some parameters in $\boldsymbol{\theta}$ that can be regarded as related or connected. In particular, we will use this *hierarchical* structure for the orbital periods and semi-amplitudes of the planets.

Within a proper Bayesian framework, all these parameters will have prior distributions associated with them, which we write as $p(N_p)$, $p(\boldsymbol{\theta})$ and $p(\boldsymbol{\alpha})$. The fact that $\boldsymbol{\alpha}$ controls for $\boldsymbol{\theta}$ means there is also a conditional prior $p(\boldsymbol{\theta} \mid \boldsymbol{\alpha})$. Assuming that the priors for N_p and $\boldsymbol{\alpha}$ are independent, and that the conditional prior for the $\boldsymbol{\theta}_i$ is independent and identically distributed and does not depend on N_p, the joint prior is factorised as

The notation $p(A \mid B)$ should read as the probability of A *given* B.

$$p(N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}_i\}_{i=1}^{N_p}) = p(N_p)\, p(\boldsymbol{\alpha} \mid N_p)\, p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_p} \mid N_p, \boldsymbol{\alpha}) \tag{2.69}$$

$$= p(N_p)\, p(\boldsymbol{\alpha}) \prod_{i=1}^{N_p} p(\boldsymbol{\theta}_i \mid \boldsymbol{\alpha}). \tag{2.70}$$

But wait, let us slow down with all the probabilities. What did that last sentence and Eqs. (2.69) and (2.70) actually mean? Independence of the priors for both N_p and $\boldsymbol{\alpha}$, and N_p and $\boldsymbol{\theta}$ means that what we believe about all, and any, of the planets' orbital parameters does not depend on how many planets there are.

Then $p(N_p)\, p(\boldsymbol{\alpha} \mid N_p) = p(N_p)\, p(\boldsymbol{\alpha})$ and $p(\boldsymbol{\theta}_i \mid N_p, \boldsymbol{\alpha}) = p(\boldsymbol{\theta}_i \mid \boldsymbol{\alpha})$.

The conditional prior for the $\boldsymbol{\theta}_i$ being independent means that the prior for the orbital period does not depend, for example, on the prior for the eccentricity. And the same for all the other orbital parameters.

Then $p(\{\boldsymbol{\theta}_i\}_{i=1}^{N_p} \mid X) = \prod_{i=1}^{N_p} p(\boldsymbol{\theta}_i \mid X)$

All of these are simplifying assumptions that make the problem

easier. We need to state these assumptions in order to write Eq. (2.70). This makes it harder to (unintentionally) hide important assumptions. Obviously, all these assumptions are subject to scrutiny.

Our goal now is to obtain properties of the posterior distribution, which is given by (this is Bayes' theorem):

$$\overbrace{p(N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}\}_{i=1}^{N_p} \mid \mathcal{D})}^{\text{posterior}} = \frac{\overbrace{p(N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}\}_{i=1}^{N_p})}^{\text{prior}} \cdot \overbrace{p(\mathcal{D} \mid N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}\}_{i=1}^{N_p})}^{\text{likelihood}}}{\underbrace{p(\mathcal{D})}_{\text{evidence}}}. \tag{2.71}$$

For the sake of clarity, we can fix, say, $N_p = 1$ and get rid of $\boldsymbol{\alpha}$ to rewrite Eq. (2.71) in more familiar terms (e.g. Gregory, 2010):

$$p(\boldsymbol{\theta} \mid \mathcal{D}) = \frac{p(\boldsymbol{\theta}) \cdot p(\mathcal{D} \mid \boldsymbol{\theta})}{p(\mathcal{D})} \tag{2.72}$$

which is a special case when we *know* the number of planets, and $\boldsymbol{\theta}$ represents the parameters for these planets.

Note how in Section 2.3.2, we chose the parameter values such that the data we measured were the most likely outcome. But we actually want the most probable parameters, given the data. This is the subtle but important difference between maximising the likelihood $p(\mathcal{D} \mid \boldsymbol{\theta})$ or the posterior $p(\boldsymbol{\theta} \mid \mathcal{D})$.

The current state-of-the-art in calculating posterior distributions of the form (2.72) is to use Markov chain Monte Carlo (MCMC) algorithms, as pioneered (in the exoplanet literature) by Ford (2003, 2005) and Gregory (2005). In principle, these same algorithms could be applied to the more general form (2.71), although in this case the MCMC sampler must be able to jump between solutions with different numbers of planets, i.e., with different numbers of parameters.

Problems like this have been studied in the statistical literature, giving rise to algorithms such as birth-death MCMC (Stephens, 2000) or reversible jump MCMC (Green, 1995). Recently, Brewer (2014) developed a new approach, within the Diffusive Nested Sampling framework of Brewer et al. (2011). A sufficiently clear explanation of Diffusive Nested Sampling would be an unwelcomed detour at this point, and is left to Appendix A. The interested reader may nevertheless jump to the Appendix and then resume at this point.

We shall assume we can obtain a number of samples[8] from a posterior distribution of the form Eq. (2.71) and continue with the assignment of each of the necessary terms.

[8] Note that one sample from Eq. (2.71) means one value for all the parameters in the model.

2.3.5 *The likelihood*

The concept of likelihood has intrigued me for some time. I mentioned it briefly in Section 2.3.2 but not in enough detail. It is such a central

concept in all of statistical inference (Bayesian or not) that it deserves to be understood. The definition from Wolfram MathWorld[9] is

> Likelihood is the hypothetical probability that an event that has already occurred would yield a specific outcome. The concept differs from that of a probability in that a probability refers to the occurrence of future events, while a likelihood refers to past events with known outcomes.

[9] Weisstein, Eric W. "Likelihood" From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Likelihood.html

Let us set N_p to 1 and think of an expression for the likelihood of a single observed RV data point, as this is an event that has already occurred. To be clear, instead of the likelihood for a RV dataset with N measurements, we want the likelihood of a single measurement.

We start by assuming that our measuring instrument is perfect, it measures the radial velocity of a star without any error (or $\sigma_1 = 0$, if you prefer). We also assume a given set of parameters $\boldsymbol{\theta}$, that completely define all aspects of the orbit of the planet. The likelihood of obtaining, at time t, any other measurement other than $v_{\rm model}(\boldsymbol{\theta}, t)$ is 0. In other words, the likelihood is a Dirac delta function, centered at the value of $v_{\rm model}(\boldsymbol{\theta}, t)$.

But there are no perfect instruments. A radial-velocity measurement made with HARPS has an associated uncertainty. Therefore, there is a finite likelihood of obtaining an outcome different than $v_{\rm model}(\boldsymbol{\theta}, t)$. How do we set the form of the likelihood now?

I could make the same argument as in the start of Sect. 2.3.2 and say that uncertainties from individual RV observations are expected to follow a normal distribution. But going from here to a Gaussian likelihood is not intuitive, because the argument is rather circular.[10]

[10] This argument goes like: assume X is normally distributed, therefore X is normally distributed.

What we do know about the measurement is the magnitude of its associated uncertainty. This is the value of the error bar associated with each RV observation. We also know[11] the mean of the likelihood distribution: it is equal to $v_{\rm model}(\boldsymbol{\theta}, t)$. These two pieces of information define, *uniquely*, a probability distribution: the Gaussian distribution. This distribution assumes the least amount of information about the measurement. So it is because we know the value of the error bar that we can assume it is normally distributed (Jaynes, 2003).[12]

[11] Although we did not have to.

[12] This argument goes like: assume we know the variance of X; then, the most conservative choice is to assume that X is normally distributed. It can be demonstrated using the principle of maximum entropy (e.g. Cover and Thomas, 2006).

A further assumption we can often make is that the N radial-velocity measurements are independent. Then the likelihood of the N measurements is the product of the likelihoods of the individual measurements. If the individual likelihood is a Gaussian

$$v_{\rm obs}(t_k) \mid N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}\}_{i=1}^{N_p} \sim \mathcal{N}\left[v_{\rm model}(\boldsymbol{\theta}, t_k); \sigma_k\right] \tag{2.73}$$

then the total likelihood is the product of N Gaussians

$$\mathcal{D} \mid N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}\}_{i=1}^{N_p} \sim \prod_{k=1}^{N} \mathcal{N}\left[v_{\rm model}(\boldsymbol{\theta}, t_k); \sigma_k\right] \tag{2.74}$$

In practice, we are often not sure that the σ_k correctly describe all sources of uncertainty in the measurements. A solution to this is to consider an additional white noise term and make its scale, s, a free parameter in the model (see Gregory, 2012, and references there in). Then, the combination of the known errors and the extra noise has a Gaussian distribution with variance $\sigma_k^2 + s^2$. The likelihood becomes

$$\mathcal{D} \mid N_p, \boldsymbol{\alpha}, \{\boldsymbol{\theta}\}_{i=1}^{N_p}, s \sim \prod_{k=1}^{N} \mathcal{N}\left[v_{\text{model}}(\boldsymbol{\theta}, t_k); \sigma_k + s\right] \tag{2.75}$$

It has been suggested (e.g. Brewer and Donovan, 2015) that a Gaussian likelihood might be overly sensitive to outliers – RV measurements that deviate substantially from the model, because of unknown effects. A t-Student distribution is often offered as an alternative, although less easily justified. Its efficiency has never been completely demonstrated, and in practical applications, we have found no compelling reasons to consider this alternative likelihood.

2.3.6 *The priors*

In any Bayesian problem, prior distributions for all parameters must be set. They can sometimes follow directly from physical arguments, but other times are simply the best way we find for encoding our state of knowledge. Table 2.1 lists the set of priors which I call "reference" priors, in the sense that these forms will be used throughout the Thesis, except when otherwise explicitly noted. The reader may notice that many of the priors have the same form as those used Brewer and Donovan (2015).

Starting with the prior for the number of planets N_p, we choose to use a uniform prior between 0 and a maximum number of planets $N_{p,\max}$. Note that this form for the prior is more adequate for small values of $N_{p,\max}$, say 3 or 4. With $N_{p,\max} = 10$, for example, we are assigning 50% probability to having more than 5 planets, which might be hard to justify a priori. In any case, a more relevant form for this prior can only come from exoplanet population studies, which address the relative abundance of single and multi-planet systems.

For the orbital periods and semi-amplitudes we will assign hierarchical priors conditional on hyperparameters μ_P, w_P and μ_K, as defined in Table 2.1. The set of these three hyperparameters corresponds to the $\boldsymbol{\alpha}$ mentioned in Section 2.3.4. They are used to set a common prior for the orbital periods, and a common prior for the semi-amplitudes, of all the planets in the system.

This reflects our belief that some parameters of one planet provide a small amount of information about some parameters of another planet. For example, the prior allows the orbital periods to cluster around a typical period μ_P while also stating that it is unlikely that

many planets have very similar or very different (relative to w_P) orbital periods. Note that within the limits allowed for w_P, the orbital periods can still be different by several orders of magnitude.

The priors for the logarithms of μ_P and μ_K are Cauchy distributions, centred on 1 yr and $1\,\mathrm{m\cdot s^{-1}}$, respectively. The heavy tails of these distributions allow for these parameters to vary over a wide range. Note that these priors are not scale invariant, meaning that one must take care to input the data in these same units.

These hierarchical priors can be easily replaced by more traditional ones (e.g. Ford and Gregory, 2007; Tuomi and Jones, 2012), with only one caveat: the priors should be the same for all the planets. This is motivated by the fact that we are building a *mixture* model, and each component of the model (each planet) should be interchangeable.

The assumption of interchangeability could of course be relaxed. If, for example, we already knew of the existence of one planet, say from a transit detection, we would be able to place a stronger prior for the orbital period of that planet. For the purposes of the Thesis though, we will not explore such cases any further.

	hyper parameters		
N_p	number of planets		$\mathcal{U}\,(0, 10)$
μ_P	median orbital period	log	$\mathcal{C}\,(5.9, 1)$
w_P	diversity orbital periods		$\mathcal{U}\,(0.1, 3)$
μ_K	mean semi-amplitude	log	$\mathcal{C}\,(0, 1)$
	planet parameters		
P	orbital period	log	$\mathcal{L}\,(\log \mu_P, w_P)$
K	semi-amplitude		$\mathcal{E}\,(\mu_K)$
e	eccentricity		$\mathcal{U}\,(0, 1)$
ϕ	orbital phase		$\mathcal{U}\,(0, 2\pi)$
ω	longitude of line-of-sight		$\mathcal{U}\,(0, 2\pi)$
s	extra white noise	log	$\mathcal{C}(0, 1)$
v_{sys}	systematic velocity		$\mathcal{U}\,(\min v_{\mathrm{obs}}, \max v_{\mathrm{obs}})$

Table 2.1: Meaning and reference prior distribution for the parameters in the RV model. In some cases we sample on the logarithm of the parameter. The distributions are:
$\mathcal{U}(\cdot,\cdot)$ – Uniform prior with lower and upper limits;
$\mathcal{C}(\cdot,\cdot)$ – Cauchy prior with location and scale (these distributions were truncated for numerical reasons);
$\mathcal{L}(\cdot,\cdot)$ – Laplace prior (often called double exponential) with location and scale;
$\mathcal{E}(\cdot)$ – Exponential prior with mean;
$\mathcal{LU}(\cdot,\cdot)$ – Log-uniform prior with lower and upper limits.

The prior for the eccentricity also merits further discussion. Having been first recognized in the analysis of stellar binary systems, there is a well-known observational bias against low eccentricities, the Lucy-Sweeney bias (Lucy and Sweeney 1971, see also Luyten 1936). This bias is actually present in any positive-definite parameter. It arises since any observational uncertainty in the measured RVs from a planetary system on a circular orbit, can only be translated in eccentricity values higher than zero.

In their MCMC analysis, Eastman et al. (2013) discuss this bias and show that the posterior distributions estimated with MCMC do not suffer from the Lucy-Sweeney bias to the same extent as classical least-squares analyses. However, as these authors also discuss, when

presenting the results of an MCMC (or any Bayesian analysis), one always needs to summarize the posterior for the eccentricity which, more often than not, will depart from a Gaussian distribution.

More recently, Zakamska et al. (2011) performed extensive simulations to assess the eccentricity bias.[13] They found that "eccentricities of planets on nearly circular orbits are preferentially overestimated, with typical bias of one to two times the median eccentricity uncertainty in a survey". They also suggested that the value of eccentricity with highest posterior probability (the mode of the marginalized posterior distributions) should be reported.

[13] See also Shen and Turner (2008).

These considerations do not relate directly to the eccentricity prior, insofar as it would be difficult to correct for a Lucy-Sweeney-type bias by imposing a given form for the prior. Another research avenue was taken by Kipping (2013), who parameterized the eccentricity distribution of known exoplanets. In principle, a Bayesian analysis can use this information about the population of exoplanets: for example, Brewer (2014), Gregory (2016), and Santerne et al. (2016b) did use the Beta distribution found by Kipping (2013) (or an approximation) as their prior for the eccentricity.

Other forms for the eccentricity prior have been proposed (e.g. Tuomi, 2012). But the extent to which the use of these priors solves the bias towards high eccentricities, or actually reflects prior knowledge, has not been adressed. For this reason, and with a simplicity argument in mind, I chose to use the uniform distribution between 0 and 1 as it is clearly the simplest choice one could consider.

Priors for the most of the remaining parameters are quite standard: uniform distributions for the angle parameters ϕ and ω, and for the systemic velocity $v_{\rm sys}$, all between sensible limits. For the extra white noise s, a Cauchy prior on the logarithm is considered which (as for μ_K) allows the parameter to vary over a very large range.

To finish this section, I present in Fig. 2.11 the histograms of 10^6 samples from the reference prior distributions. For this example, a model with $N_{p,\max} = 3$ was considered, and the limits for $v_{\rm sys}$ correspond to the RV peak-to-peak limits of HD 111515. The histogram of the orbital period prior samples (and only that one) also contains a representation (as shaded regions) of the conditional priors $p(P \mid N_p = i)$ for $i = 1, 2, 3$. Note that these conditional priors all have the same shape, which results from our assumption of independence of the priors for N_p and $\boldsymbol{\theta}$.

2.3.7 *The evidence*

The last term in Eq. (2.71) that we still need to discuss is the evidence, $p(\mathcal{D})$. Because the posterior is a probability distribution (and thus integrates to unity over its support) and because the evidence does

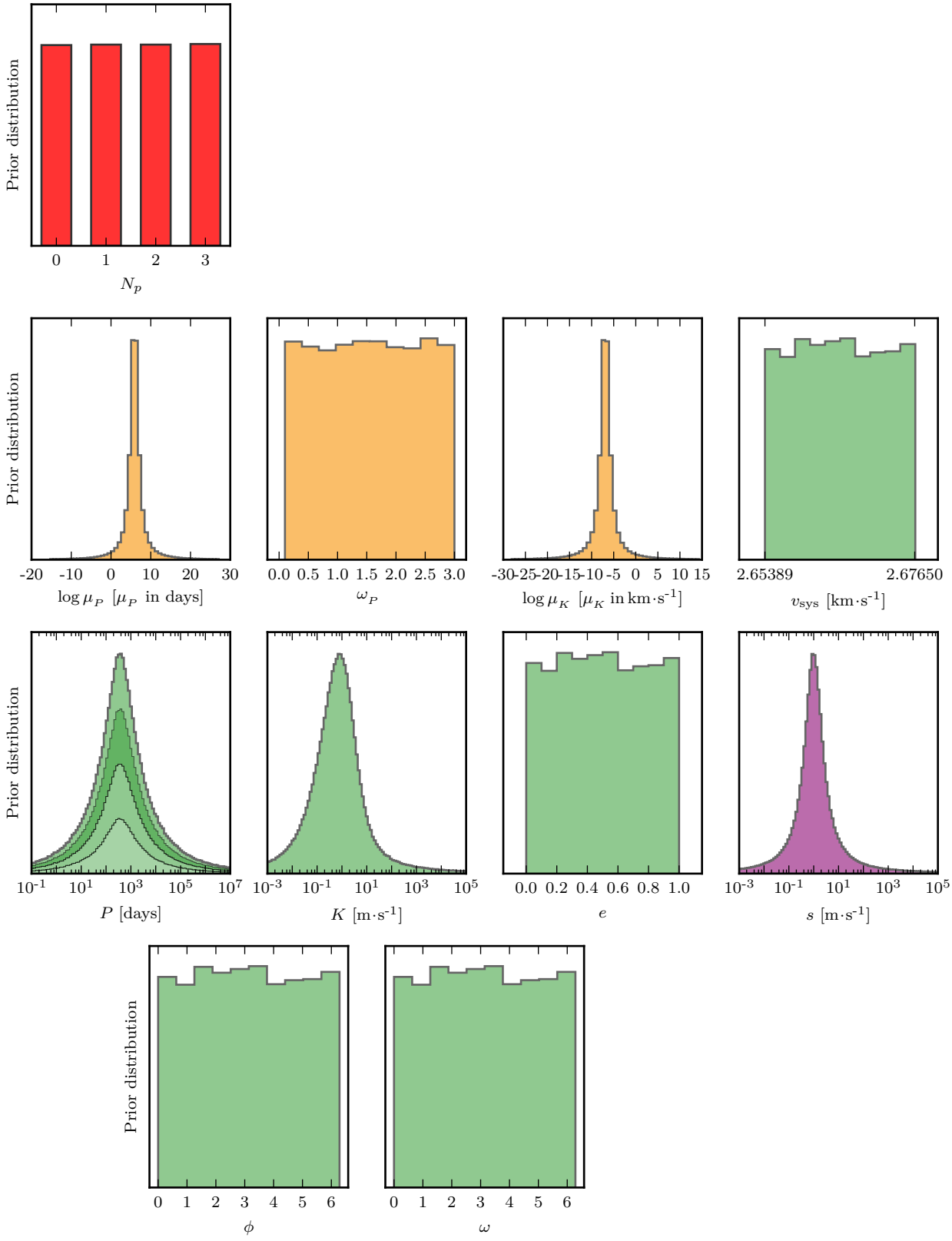


Figure 2.11: Histograms of 10^6 samples from the reference prior distributions. Note the different units of the parameters and that when the prior is specified for the logarithm of the parameter, the unit is also mentioned.

not depend on N_p, $\boldsymbol{\theta}$ or $\boldsymbol{\alpha}$, we can rewrite Eq. (2.71) as

$$\int_{N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p}} p(\mathcal{D}) \cdot p(N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p} \,|\, \mathcal{D}) \;=\; \int_{N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p}} p(N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p}) \cdot p(\mathcal{D} \,|\, N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p})$$

and perform the integration on the left-hand side

$$p(\mathcal{D}) \;=\; \int_{N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p}} p(N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p}) \cdot p(\mathcal{D} \,|\, N_p,\boldsymbol{\alpha},s,\{\boldsymbol{\theta}\}_{i=1}^{N_p}) \qquad (2.76)$$

The evidence is then the average of the likelihood over the prior. Its value is larger for a model where more of the parameter space is likely and is smaller for a model with large areas in its parameter space having low likelihood values. This is what is meant when one says the evidence automatically implements Occam's razor: a simpler model with a more compact parameter space will have a larger evidence than a more complicated one, unless the latter is significantly better at explaining the data.

Given two different models A and B, Bayesian model comparison is interested in comparing the posterior probabilities of the models, $p(A \,|\, \mathcal{D})$ and $p(B \,|\, \mathcal{D})$. Using Bayes' theorem, and assuming equal prior probabilities for both models, we compute the *Bayes factor*, as

$$\ln \mathcal{B}_{A,B} = \ln \left[\frac{p(A \,|\, \mathcal{D})}{p(B \,|\, \mathcal{D})} \right] = \ln \left[\frac{p(\mathcal{D})_A}{p(\mathcal{D})_B} \, \frac{p(A)}{p(B)} \right] = \ln \left[\frac{p(\mathcal{D})_A}{p(\mathcal{D})_B} \right] \qquad (2.77)$$

where $p(\mathcal{D})_A$ is the evidence of model A and $p(\mathcal{D})_B$ of model B. Therefore, when both models are equally probable a priori, the ratio of posterior probabilities is equal to the ratio of evidences.

$\mathcal{B}_{A,B}$	$\ln \mathcal{B}_{A,B}$	$p(A \,\|\, \mathcal{D})$	descriptive label
> 150	> 5	> 0.993	decisive evidence for model A
20 to 150	3 to 5	0.952 to 0.993	strong evidence for model A
3 to 20	1 to 3	0.75 to 0.952	moderate evidence for model A
1 to 3	0 to 1	0.5 to 0.75	anecdotal evidence for model A
1	0	0.5	no evidence for either model
1/3 to 1	-1 to 0	0.25 to 0.5	anecdotal evidence for model B
1/20 to 1/3	-3 to -1	0.048 to 0.25	moderate evidence for model B
1/150 to 1/20	-5 to -3	0.007 to 0.048	strong evidence for model B
$< 1/150$	< -5	< 0.007	decisive evidence for model B

Table 2.2: Scale for the interpretation of the Bayes factor. Adapted from the scale first proposed by Jeffreys (1961), which was later simplified in Kass and Raftery (1995). The original labels were modified as in Lee and Wagenmakers (2013). Note that the values for the probability $p(A \,|\, \mathcal{D})$ assume the following: $p(A) + p(B) = p(A \,|\, \mathcal{D}) + p(B \,|\, \mathcal{D}) = 1$, meaning that only models A and B are possible.

Although the Bayes factor $\mathcal{B}$ is a continuous measure of evidence for (or against) a model, we would like to have a set of labels or categories, to provide guidelines on how to interpret its numerical value. It is clear that any such labelling will introduce (more or less arbitrary) cutoffs and be context dependent. This said, a number of 'scales' were introduced for interpretation of Bayes factors, the most

common system being that of Jeffreys (1961), later modified by Kass and Raftery (1995). This scale is presented in Table 2.2.

The reader will note that taking the logarithm of $\mathcal{B}$ has the advantage that the scale above and below $\mathcal{B} = 1$ is the same, and only the change of sign reflects preference for either model A or model B. This also makes $\ln \mathcal{B}$ more suited for visualisation purposes. One may note an interesting attempt to reconcile Jeffreys' scale for Bayes factors with the classical interpretation of model comparison based on p-values, as discussed in Efron and Gous (2001).

In our model, the value of the evidence is not actually used for model comparison. It will be the posterior distribution for N_p that interests us, as this distribution already answers the question of how many planets are detected in the data. In any case, (Diffusive) Nested Sampling calculates the evidence together with the posterior for all the parameters (see Appendix A) and the value can be used in the future to compare other models for the same data. Of course, we have the possibility to fix N_p to, e.g., 0 and 1, and use the evidence value to compare these two models.

We can also calculate the *information* $\mathcal{H}$, or Kullback–Leibler divergence (Kullback and Leibler, 1951), from the prior to the posterior,[14]

$$\mathcal{H} = \int_{N_p, \boldsymbol{\alpha}, s, \{\boldsymbol{\theta}\}_{i=1}^{N_p}} p(N_p, \boldsymbol{\alpha}, s, \{\boldsymbol{\theta}\}_{i=1}^{N_p} \,|\, \mathcal{D}) \cdot \ln \left[\frac{p(N_p, \boldsymbol{\alpha}, s, \{\boldsymbol{\theta}\}_{i=1}^{N_p} \,|\, \mathcal{D})}{p(N_p, \boldsymbol{\alpha}, s, \{\boldsymbol{\theta}\}_{i=1}^{N_p})} \right] \quad (2.78)$$

which quantifies how much we learned about the parameters. Intuitively, the information is the number of times the prior distribution had to be compressed by a factor of e to obtain the posterior.

I will present values of $\mathcal{H}$ in natural units, 'nats'. A value of $\mathcal{H} = 10$ nats means the posterior occupies e^{-10} times the volume of the prior. When analysing two datasets with the same model (so that the prior volume is the same), this quantity can be used to measure how much information is contained in each dataset, and therefore how much each dataset is constraining the parameters.

[14] For texts on *information theory* see for example Cover and Thomas (2006), MacKay (2003) and the foundational Shannon and Weaver (1998).

2.3.8 *The full model*

Given expressions for the likelihood and priors, our inference problem is essentially solved and we can use the DNS algorithm to sample the posterior distribution and calculate the value of the evidence.

In total we have $5 \times N_p + 6$ parameters, the orbital parameters for each of the N_p planets, the hierarchical parameters μ_P, ω_P and μ_K, the number of planets N_p, the extra white noise term s, and the systemic velocity $v_{\rm sys}$ of the star (see Table 2.1). The full model is represented

in the probabilistic graphical model (PGM) of Fig. 2.12. This figure shows all the parameters of the model and how they relate to each other and to the data. Each *node* (a parameter with one circle around it) has associated prior and posterior distributions and the arrows indicate how the parameters are logically related.

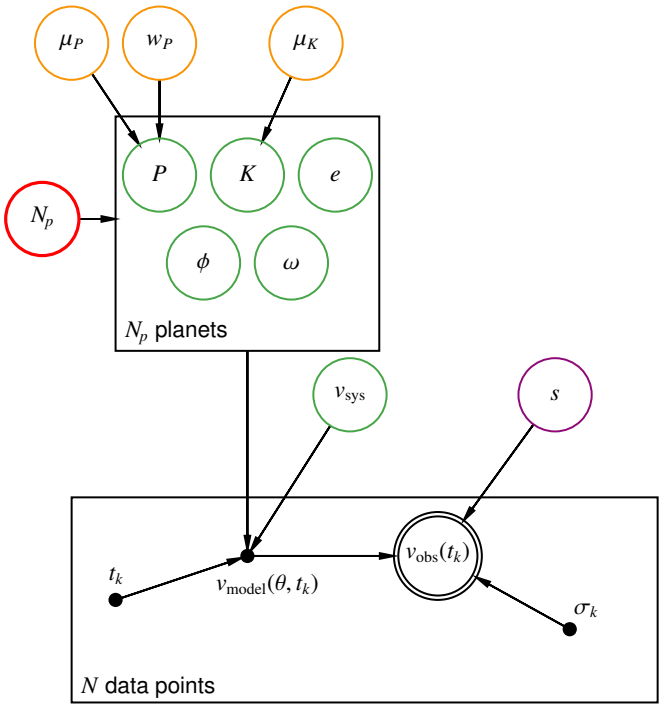


Figure 2.12: Representation of the relations between parameters and observations in the RV model, as a probabilistic graphical model (PGM). An arrow between two *nodes* indicates the direction of conditional dependence. The circled nodes are the parameters of the model, whose joint distribution is sampled by the MCMC algorithm. These parameters therefore have prior and posterior distributions associated with them. The double circled node $v_{obs}(t_k)$ represents the observed RVs. The filled nodes represent deterministic variables: if these variables have parent nodes – $v_{model}(\boldsymbol{\theta}, t_k)$ – they are given by a deterministic function of those parents; if there are no parents – t_k and σ_k – they are assumed given and thus fixed. The variables inside boxes are repeated a number of times shown in the bottom left corner of the box. Adapted from Faria et al. (2016a).

For example, μ_P and ω_P affect the orbital periods P_i of each of the N_p planets. Almost all parameters (except s) affect the calculation of the Keplerian function $v_{model}(\boldsymbol{\theta}, t)$. In turn, s only affects the likelihood which, in the PGM, is represented by an *observed node*, the RV measurements $v_{obs}(t_k)$. This observed node has its own distribution, the Gaussian likelihood. Both the times t_k and the RV uncertainties σ_k are assumed given without any error.

The fact that both parameters and data have distributions associated with them is one of the most important aspects of Bayesian probability. Once this is grasped, nothing stops from using the terms 'unobserved' and 'observed' variables to replace 'parameters' and 'data'. Then, as shocking as it may seem, the likelihood actually becomes a prior.

* ◯ *

The posterior distribution for N_p is one of the main outputs of this method. But to actually decide on what is the number of detected planets, we have to come up with an *utility function* to analyse this distribution. Simply taking the mean or the median value of the distribution does not adequately convey its somewhat special character.

Based on the scale for Bayes factors (Table 2.2), we can instead set a probability ratio threshold for models with consecutive numbers of planets. We assert that in order to claim a detection of N_p planets, the probability of N_p should be (at least) 150 times greater than the probability of $N_p - 1$. This criterion, which I will call the *posterior*

ratio criterion, requires "decisive" evidence for detecting a planet, thus considering false positives to be much worse than false negatives.

From another perspective, if we had fixed N_p to consecutive values $0, 1, 2, \ldots$, and calculated the evidence for each model in different runs, the criterion above requires that the Bayes factor between the consecutive models should be higher than 150 to claim a detection. This is in line with, or slightly less conservative than, similar criteria from the literature (Tuomi et al., 2014b).

2.4 TESTS WITH KNOWN SYSTEMS

The probabilistic model presented above is mostly based on the work of Brewer and Donovan (2015), who already presented some examples of its application. In this Section, we will perform some further tests, with data from simulated and real planetary systems.

We will analyse: *(i)* the two simulated datasets created by Balan and Lahav (2009), *(ii)* a situation with more free parameters than data points *(iii)* the ELODIE RV data for 14 Her, and *(iv)* the HARPS RVs of HD 10180. With these analyses, we will try to get a better feeling of the basic procedure for planet detection with our probabilistic model.

2.4.1 *Simulated data*

To test their `EXOFIT` software, Balan and Lahav (2009) created two sets of simulated RV data (with 34 points), containing 1 and 2 planets. A plot of the radial velocities and the parameters used to create these data are shown in Fig. 2.13.

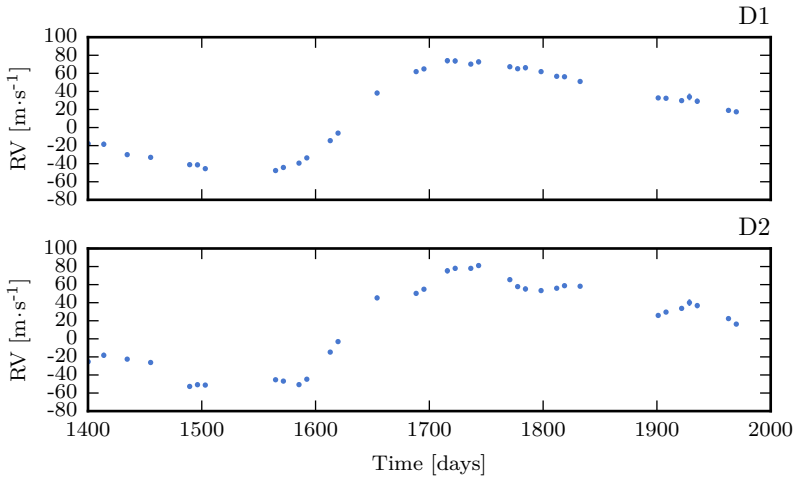


Figure 2.13: Simulated data from Balan and Lahav (2009) with 1 ($D1$) and 2 ($D2$) planets. The following table shows the parameters of the injected signals.

Parameter	$D1$	$D2$	
v_{sys}	12	12	m·s^{-1}
P_1	700	700	days
K_1	60	60	m·s^{-1}
e_1	0.38	0.38	
ω_1	3.1	3.1	
P_2	–	100	days
K_2	–	10	m·s^{-1}
e_2	–	0.18	
ω_2	–	1.1	

I analysed both datasets ($D1$ and $D2$) with the model presented in the previous section, and the maximum number of planets set to $N_{p,\max} = 4$. A total of 15670 and 20730 posterior samples were

obtained for $D1$ and $D2$, respectively. This took about 20 minutes each on a common desktop PC.

The posterior distribution for N_p is shown in Fig. 2.14, for the dataset containing 1 planet at the top, and for the one containing 2 planets at the bottom. The posterior modes correspond, in both cases, to the correct number of planets. Both distributions in Fig. 2.14 also show the ratios of consecutive probabilities $p(N_p)/p(N_p - 1)$. Since there are no posterior samples with $N_p = 0$ for $D1$ and with $N_p = 0, 1$ for $D2$, the corresponding ratios were set to infinity. This shows that, applying the posterior ratio criterion also recovers the correct number of planets in each case.

In fact, it is interesting to discuss why there are no posterior samples with $N_p = 0$ for the case of $D1$, for example. This does not mean necessarily that the probability of $N_p = 0$ is exactly zero. It only means that the ratio $p(N_p = 1)/p(N_p = 0)$ is higher than the total number of posterior samples we have obtained (in this case, 15670). Therefore, representing a probability ratio with 'inf', as in Fig. 2.14, means only a value larger than the total number of samples.

In our model, all the orbital parameters are recovered at the same time as N_p, and we can use the posterior samples to evaluate the best models in data-space (RV versus time). Fig. 2.15 shows the main results from the analysis of $D1$ and $D2$. The posterior distributions for the orbital periods, semi-amplitudes and eccentricities are shown, together with the parameter estimates they provide.

The figure also shows the best-fitting curves, obtained with the median values of the posterior distributions, and the residuals. The rms scatter of the residuals is very similar to the Gaussian noise inserted in the data by Balan and Lahav (2009).

This first test is simple (the mock data from Balan and Lahav (2009) is admitedly very clear) but provides initial clues that the method is performing well. All the orbital parameters recovered here, agree within uncertainties to the values used to simulate the data. Also, the correct number of planets is recovered.

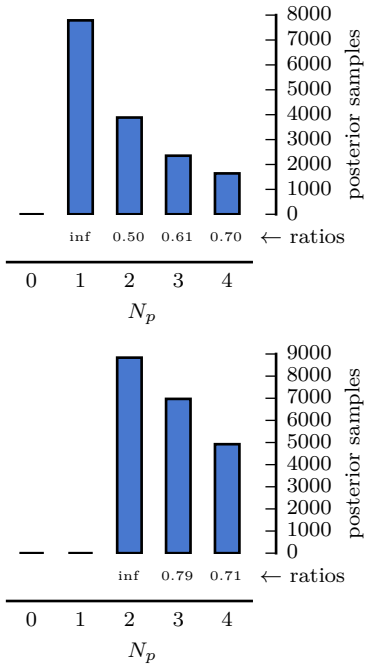


Figure 2.14: Posterior distributions for N_p for both datasets from Balan and Lahav (2009). *Top*: $D1$; *bottom*: $D2$.

2.4.2 *A situation with many parameters*

In this brief section, I try to test the behaviour of the method in a pathological situation. This analysis tries to answer a key question the reader might have: what happens if we have more free parameters than data points?

We use HARPS data for HD 128340, the star from the metal-poor sample (see Chapter 4) for which we have the fewest observations ($N = 5$). The maximum number of planets is set to $N_{p,\max} = 20$, with a uniform prior for N_p. This means that a total of $20 \times 5 + 6 = 106$ parameters are free in the model.

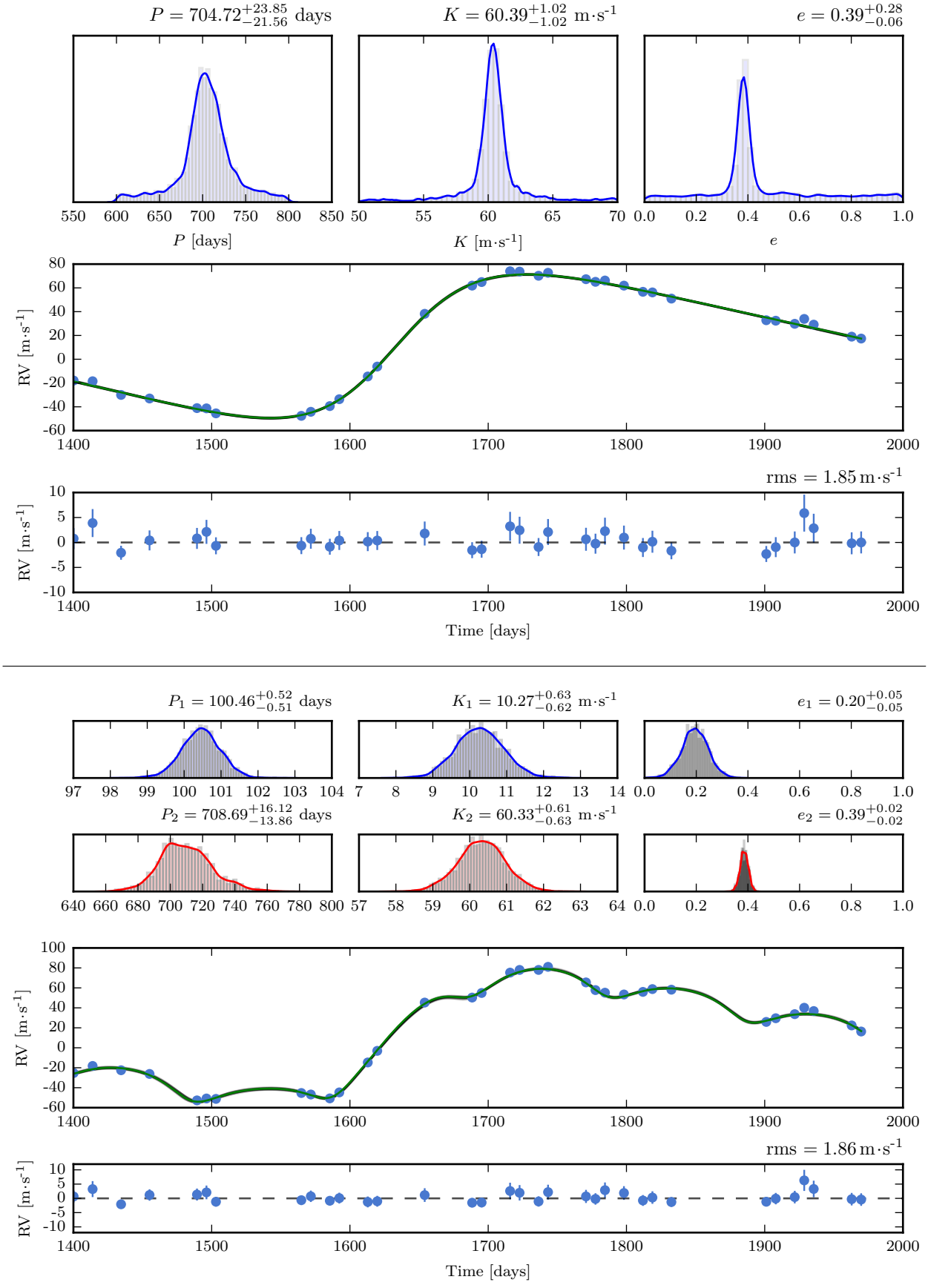


Figure 2.15: Results from the analysis of datasets *D*1 and *D*2. See text for details.

With 18308 samples, we obtain the posterior for N_p shown in Fig. 2.16. The results confirm that no planets are confidently detected in this dataset, since the mode of the distribution is $N_p = 0$. We find that, in the absence of information in the data, the resulting posterior distributions for the orbital periods, semi-amplitudes and eccentricities (of all 20 Keplerians together) are simply the same as the input priors.

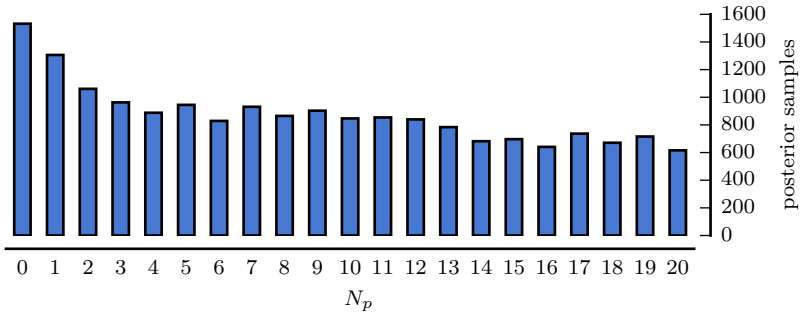


Figure 2.16: Posterior for N_p from the analysis of 5 RV measurements of HD 128340.

This test is clearly unrealistic, but nevertheless quite insightful. In nearly all analysis of RV surveys, stars with fewer than a threshold number of observations are excluded from the search for planetary signals. The threshold is usually set at $N = 6$ or, even more arbitrarily, at higher values, because $5N + 1$ parameters (orbital parameters and offset) are "needed" to fit N observations.

In our model this threshold is not necessary; the essential problem with more parameters than observations is that there is potentially no information in the data to constrain all the parameters. But if this is the case, the model itself should tell us this, a posteriori. It should not need this information as an input.

In fact, it is interesting to look at the value for the information. For this analysis of HD 128340, I obtain $\mathcal{H} = 1.3$ nats.This means that the posterior distribution occupies about 27% of the volume of the prior and since the latter was very large, it means the data does not constrain most of the parameters. Part of the compression comes from N_p, for which the prior was uniform and the posterior is somewhat constrained (Fig. 2.16) and also from the constrain to $v_{\rm sys}$.

2.4.3 *14 Herculis*

The same ELODIE data presented in Section 2.3.2 (see Fig. 2.10), is analysed here with the new model. The planet 14 Her b was first announced by Mayor (1998, oral presentation) and Marcy et al. (2000). Its updated orbital parameters were determined by Udry et al. (2000) and the detection was later confirmed with Keck/HIRES data by Butler et al. (2003).

Naef et al. (2004) analysed the same 119 RV measurements and found very large residuals from a 1-planet fit, and evidence for a linear drift. More recently, Wittenmyer et al. (2007) combined the RVs from the McDonald Observatory, Keck/HIRES and ELODIE, detecting a longer period signal, probably caused by a 2 $M_{\rm Jup}$ planet.

Seting $N_{p,\max} = 3$, about 11 000 samples from the joint posterior distribution we obtained, considering the 119 ELODIE observations. The posterior for N_p reveals the confident detection of 2 planets (Fig. 2.17). The ratio of probabilities suggests moderate evidence for a model with 3 Keplerians, which is not enough to pass the posterior ratio criterion.

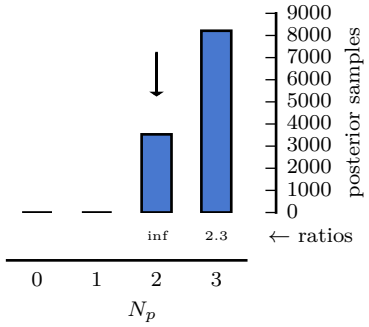


Figure 2.17: Posterior distribution for N_p from the analysis of 14 Her.

For the orbital periods, the posterior distribution shows two main peaks, around 1700, and 4500 days, corresponding to the two detected planet signals. The maximum likelihood solution, from those with $N_p = 2$, is shown in Fig. 2.18, together with the phased solutions for the two planets. The middle panel of the figure shows the residuals after subtracting the planetary signals (note the different units).

From the analysis of the posterior distribution, we can obtain estimates for the orbital parameters of the planets. For 14 Her b, we obtain a $4.49^{+0.16}_{-0.58}$ $M_{\rm Jup}$ planet orbiting at 1730.87 ± 15.16 days and with an eccentricity 0.35 ± 0.02.

On the other hand, for 14 Her c, the ELODIE observations do not cover a full period and the orbital parameters are much less constrained. The extra observations used by Wittenmyer et al. (2007) do cover the whole longer period. Our estimate for the orbital period therefore has a much higher uncertainty: $P = 5087^{+2449}_{-1162}$ days. For the planet mass we estimate $m_p \sin i = 1.92^{+1.47}_{-0.57}$ $M_{\rm Jup}$.

These results are in general agreement with those of Wittenmyer et al. (2007), but a more careful comparison is out of the scope of this section. Our main conclusion is the significant detection of two planetary companions for 14 Her. It is important to mention that these results are obtained directly from the joint posterior distributions (for N_p and for the orbital parameters).

2.4.4 HD 10180

The planetary system around HD 10180 is one of the richest ever discovered with the RV method. Lovis et al. (2011b) announced the discovery of up to 7 planets around this solar-type star, and claims of up to 9 planets followed soon after (Tuomi, 2012). With the same data, Feroz et al. (2011b) only found convincing evidence for 6 planets in the system. Here I present the analysis of the 190 RV observations obtained with HARPS, between November 2003 and June 2009.

We start by assuming the reference uniform prior for N_p, between 0 and 10 planets. In this case, we can claim that *a priori* we are

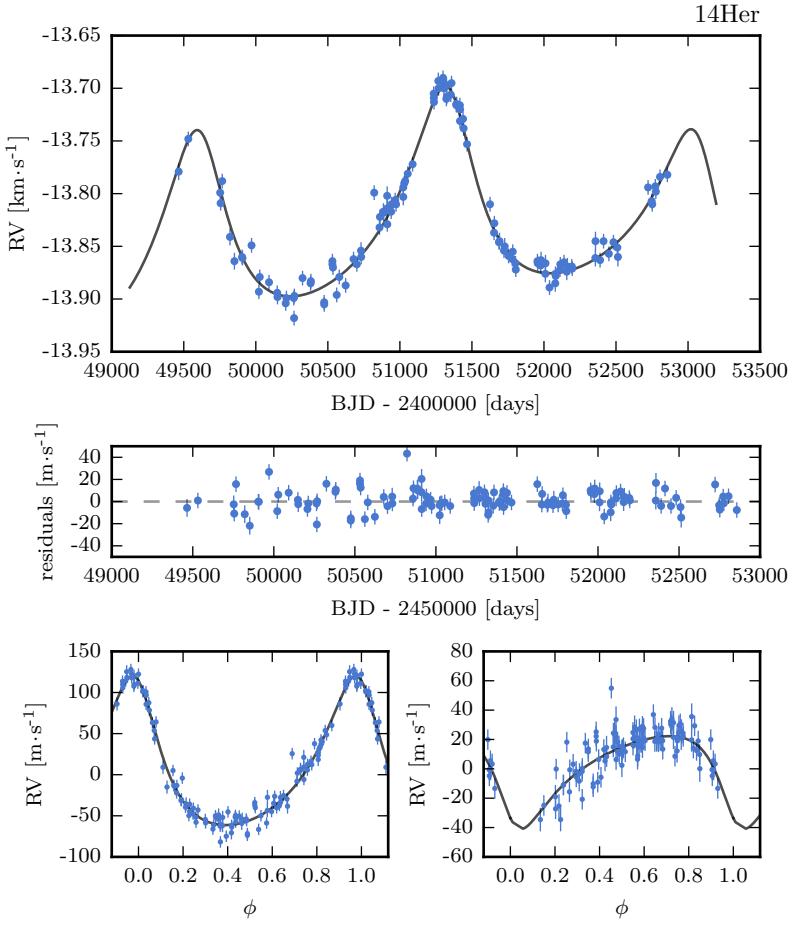


Figure 2.18: Maximum likelihood 2-planet solution for 14 Her, when considering the 119 ELODIE observations. The middle panel shows the residuals after subtracting this solution. The bottom panels show the phase curves for each planet, after subtraction of the Keplerian signal from the other.

expecting to find a large number of planets and so this prior is acceptable. Nevertheless, we can only make this claim because we know about the previous analyses (some of which used the same data) and so this is not strictly prior information.

After obtaining 17002 samples, the posterior distribution for N_p is shown in Fig. 2.19. The posterior rises continuosly from $N_p = 6$ until the border of the prior, $N_p = 10$. Going from 6 to 7 planets corresponds to an odds ratio of 1.86, meaning that the number of confidently detected planets in this dataset is only $N_p = 6$.

This is the conclusion we draw from the posterior, even if the mode is different than $N_p = 6$ and there are actually very few samples with that number of planets.

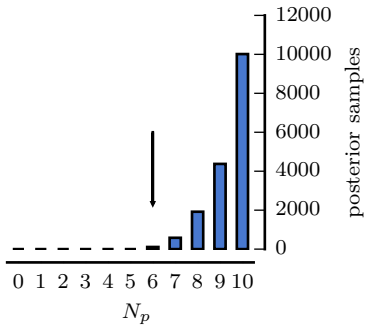


Figure 2.19: Posterior distribution for the number of planets; analysis of HD 10180 HARPS data with uniform prior for N_p.

Figure 2.19 can lead us to many questions. Is the number of samples with $N_p = 6$ enough to well characterize the orbital parameters? Does the posterior distribution have a peak at some value of N_p, or does it keep always increasing? Is our prior reasonable?

The answer to the first question is likely yes. Because we are dealing with a mixture model, every sample from the posterior for a model with $N_p = 7$, for example, is a valid sample for the model with $N_p = 6$. So the six most probable values for the orbital period, semi-amplitude, eccentricity, etc, can be assigned to the six detected planets. Or, if we are not satisfied with this solution, we can always fix N_p to 6 and just run the model again.

To try to answer the other questions we can do a couple of tests: first, we increase $N_{p,\max}$ to, say, 40. This way, we can check what happens when the model starts to be undetermined[15] and see if the posterior indeed has a mode. Then we can change to a uniform prior between 0 and 5. Since we are almost sure that there are *at least* 6 planets in this system, it is interesting to explore this more restricitive prior. The posteriors from these two tests are in Fig. 2.20.

[15] We are analysing 190 measurements, with $40 \times 5 + 6 = 206$ parameters.

With the uniform prior over a broader range, we can see that the posterior peaks at $N_p = 19$ and falls off for higher values. Interestingly, the posterior ratio criterion still gives $N_p = 6$ as the number of confidently detected planets, even if the actual probability of $N_p = 6$ is small. Another interesting aspect to note is the fact that the ratio $p(N_p = 19)/p(N_p = 6)$ is actually very large ($\sim$ 459) and well above the threshold for decisive evidence. Even if counter-intuitive at first, this is not an issue, since the important aspect when deciding on how many planets are detected is the ratio of *consecutive* probabilities.

To understand this very important point in another way, we could explore the posterior, starting from the mode and ask "is there enough evidence to choose $N_p = 19$ over $N_p = 18$?" to which the answer will be "No". By repeating this process, we would end up asking "is there enough evidence to choose $N_p = 6$ over $N_p = 5$?" to which the

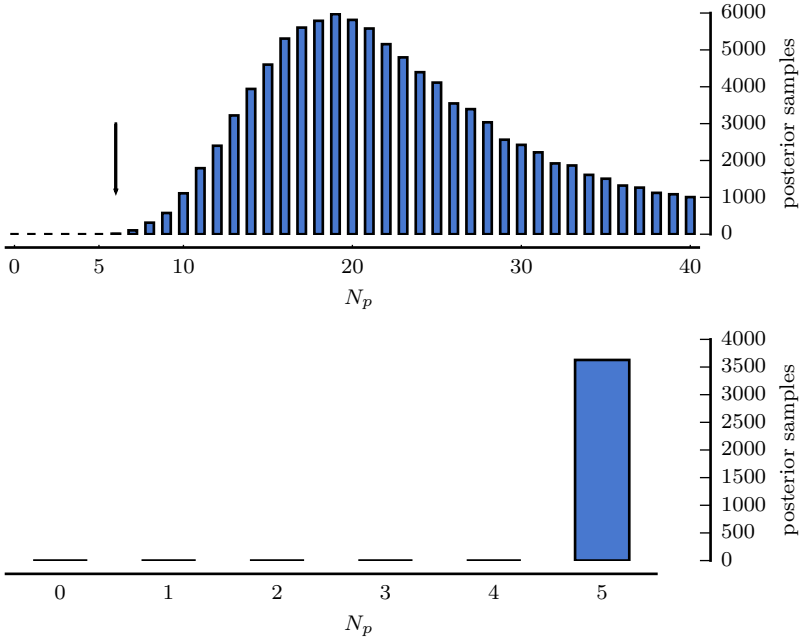


Figure 2.20: Posterior distributions for N_p obtained with two different priors: *top*: $\mathcal{U}(0,40)$, *bottom*: $\mathcal{U}(0,5)$. In the top panel, the arrow shows the value of N_p resulting from the posterior ratio criterion. In the bottom panel this value is $N_p = 5$.

answer is "Yes."

The results with the more restrictive prior are arguably even more interesting. The posterior for N_p is absolutely decisive for the value $N_p = 5$. If *a priori* we expected between 0 and 5 planets, after having observed the data we are sure that there are 5 planets.

Looking at the posteriors for some of the orbital parameters, another clear result emerges: we have detected the 5 highest amplitude signals out of the (at least) 6 present in the data. Fig. 2.21 displays this result, with the red triangles representing the parameters determined by Lovis et al. (2011b) for 6 of their reported 7 planets.

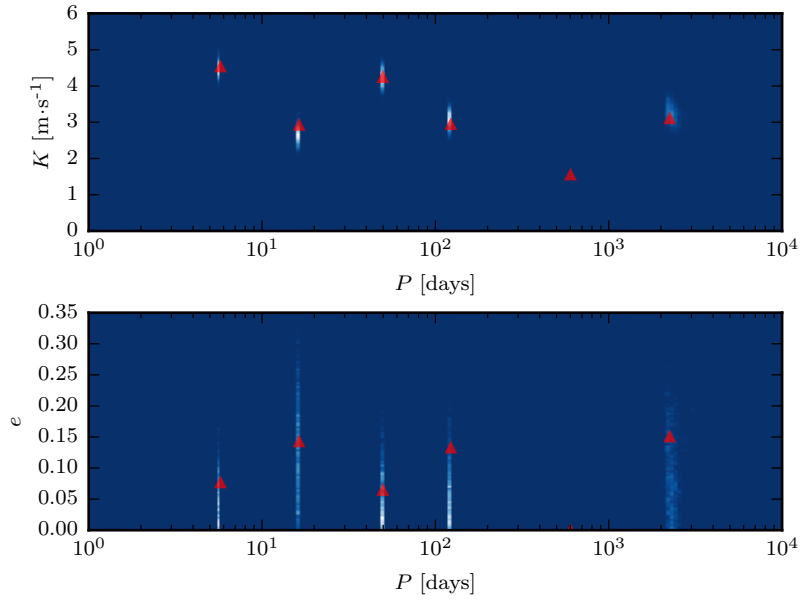


Figure 2.21: Joint posterior distributions for the semi-amplitudes, eccentricities and orbital periods of HD 10180, obtained with a uniform prior for N_p between 0 and 5. Red triangles show the same parameters obtained by Lovis et al. (2011b), for the six outer planets of their 7-planet Keplerian fit. Note that the eccentricity of HD 10180 g, at 600 days, was fixed to 0 by Lovis et al. (2011b).

This result aligns with our intuition that the highest amplitude signals should be the "easiest" to detect, and also shows how the method responds to a situation where the prior (at least for N_p) is too restrictive. In any case, these results with a $\mathcal{U}(0,5)$ prior for N_p would prompt us to increase $N_{p,\max}$ and test for additional planets.

3

Effects of stellar activity

The number of exoplanets detected with the RV method would certainly be higher if exoplanets were the only causes of the signals found in stellar radial velocity timeseries. In fact, the stars themselves are not quiet. The many different kinds of *stellar activity*[1] also induce variations in the measured radial velocities.

Different physical phenomena in the stellar photospheres produce signals with different timescales, going from a few minutes to several years. In addition, the RV manifestations of these phenomena are often periodic, or nearly so, and can quite easily mimic the signals caused by planets, which have comparable amplitudes. The detection of low-mass planets, even in moderately active stars, requires an understanding of how to deal with these contaminations.

In this Chapter, I briefly review the physical processes behind these sources of noise and the resulting signals imprinted in the RVs (Section 3.1). Some of the current strategies to deal with or correct for activity-induced signals are then explained (Section 3.2). In Section 3.3, I present a new model for the analysis of RV datasets of active stars, which uses Gaussian processes as correlated noise models. A deliberately more didactic introduction to Gaussian processes is also included in this Section. Finally, in Section 3.4, the results of applying this model to the case of CoRoT-7 are detailed. These results were first published in Faria et al. (2016a).

[1] Note that throughout this Chapter, I may fall into the trap of using the term stellar activity to denote a variety of different phenomena. Not all of these phenomena have the same origins, and in particular not all can be considered stellar *magnetic* activity, so the term might be considered abusive. The reader is free to think of stellar activity, as I use it here, as a global term to mean stellar pulsations, stellar granulation and stellar magnetic activity.

3.1 THE ACTIVE STAR AND ITS RADIAL-VELOCITY SIGNATURES

In the Sun, the cyclic regeneration of a large-scale magnetic field gives rise to the different types of solar activity (e.g. Charbonneau, 2010). This regeneration is likely caused by a hydromagnetic dynamo process, whose source is the inductive action of fluid motions in the solar interior. The radiative zone in the interior of the Sun is separated from the outer convective layer (see Fig. 3.1) by the *tachocline* (Spiegel and Zahn, 1992). In the tachocline, the transition from the radiative zone, which rotates as a solid body, to the differentially rotating convection zone, creates strong shearing forces, which fuel the magnetic dynamo (see e.g. Tobias, 2002)

Figure 3.1: Internal structure of stars with different masses. Solar-mass stars have radiative cores and convective envelopes, in massive stars the situation is reversed, and low mass stars are fully convective.

Stellar activity in solar-like stars is thought to be caused by the same dynamo effect (Noyes, 1986; Schrijver and Zwaan, 2008). In less massive, fully convective stars, other types of dynamos can also give rise to magnetic activity (see for example Chabrier and Küker, 2006; Route, 2016; Houdebine et al., 2017). On the other hand, stars that are fully radiative usually show much weaker magnetic fields, if any (Walder et al., 2012), with the notable exception of chemically peculiar Ap stars (e.g. Cunha, 2009).

The observational signatures of stellar activity and other phenomena taking place at the stellar surfaces, are commonly presented in terms of their associated timescales, since their spatial scales are usually inaccessible in stars other than the Sun. In this regard, different phenomena can be observed, ranging from a few minutes to several years, and are divided in four main contributions: solar-like oscillations, granulation, short-term activity caused by surface active regions and long-term activity from variations in the overall stellar magnetic field. Note that the first two are *not* caused by the stellar magnetic field, even if they are often included in the same "activity bag".

Fig. 3.2 shows a pictorial representation of the timescales associated with active regions, granulation and oscillations (see for example Aigrain et al., 2004; Dumusque et al., 2011c). The power spectrum[2] is dominated at lower frequencies, on the order of the stellar rotation period, by the signals induced by active regions. At timescales of a few hours, granulation impacts another power-law contribution (Harvey, 1985; Karoff, 2008; Kallinger et al., 2014), and at higher frequencies, oscillations can be characterized by a Gaussian "bump" of excess power (e.g. Kjeldsen et al., 2005). Note that, in principle, the contribution of each component can be essentially averaged-out to white noise by binning on appropriate timescales.

[2] See for example Stoica and Moses (2005) for a definition.

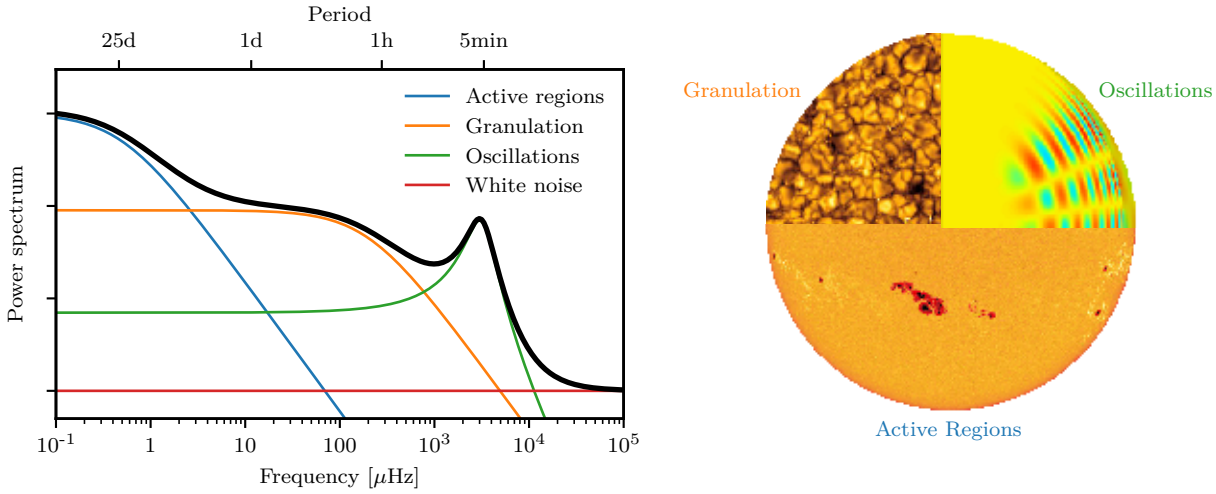


Figure 3.2: Illustration of the timescales associated to stellar activity, represented by the power spectrum of the different activity-induced signals.

All these activity phenomena cause variations in the measured stellar RV, to different extents. Possibly the easiest to understand perturbation affecting the RVs, arises from the temperature differences between active regions and the surrounding stellar surface – what is usually called the *flux effect* (e.g. Dumusque et al., 2014a). This effect can result in a RV variation of many $m \cdot s^{-1}$, depending on the size of the active regions and the projected rotational velocity of the star (Saar and Donahue, 1997)

The observed stellar spectrum is the sum of the spectra emitted by individual surface elements on the stellar surface. For a star seen equator on,[3] half of the visible hemisphere is moving towards us, and those surface elements are blueshifted. The other half is moving away (with the same velocity) and thus is seen redshifted. For an homogeneous star, the net effect is null and the star shows no RV variation as it rotates.

[3] And having a rotational velocity $v \sin i \neq 0$.

A spot on the stellar surface has a smaller flux when compared to the average stellar surface, because its temperature is lower (Solanki, 2003; Thomas and Weiss, 2012). The presence of the spot and the fact that it rotates together with the star (see Fig. 3.3), creates a time-variable photometric and spectroscopic signature. When the spot is on the approaching (resp. receding) half-hemisphere, it blocks blueshifted (resp. redshifted) light. The net RV effect is no longer zero: there is a rotational modulation of the observed RV. As the star rotates, there is a perturbation to the CCF which changes its centre (Fig. 3.3). This effect is at the basis of the Doppler imaging technique (see e.g. Donati et al., 1997; Barnes et al., 2017).

It is clear that this effect will be present for more than one spot too, as long as there is an inhomogeneity in the spot distribution over the stellar surface. It is also easy to understand why the magnitude of the effect depends on the rotational velocity of the star, since that controls how blueshifted and redshifted each half-hemisphere is. In addition, spots grow and decay and eventually disappear from the stellar surface (e.g. Thomas and Weiss, 2012), with new ones appearing at a different disk position (and thus a different "phase"). This introduces a time variation in both the photometric and spectroscopic signals, which would otherwise be periodic.

In its turn, the RV effect caused by stellar granulation, is a result of convective motions in the stellar atmospheres (Nordlund et al., 2009). In the photosphere, hot (thus bright) and rising (thus blueshifted) convective elements contribute more photons than the cooler (darker) and sinking (redshifted) gas, which causes a net blueshift of the absorption lines (e.g. Dravins, 1982; Lindegren and Dravins, 2003).

This blueshift is of the order of $300\,m \cdot s^{-1}$, although it depends on the line depth (Gray, 2009; Meunier et al., 2010a). Variations

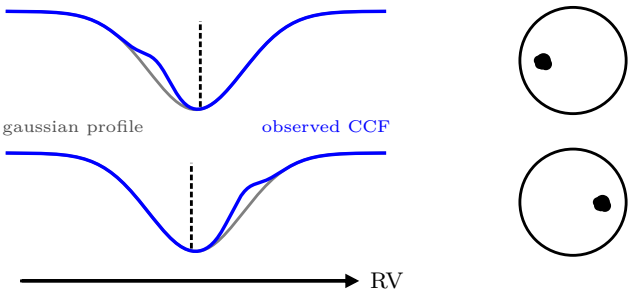


Figure 3.3: Effect of a stellar spot on the observed CCF, as the star rotates. The spot causes a distortion in the CCF, which shifts its center of mass and therefore the measured RV. Note that this effect is extremely small and that the CCF is represented here with unrealistic deformations for visual clarity.

over this net RV effect can arise on timescales from minutes to a few hours (Dumusque et al., 2011c), and reach amplitudes of around $2\,\mathrm{m{\cdot}s^{-1}}$ (Meunier et al., 2010a). Stellar granulation also has a clear photometric signature, which is now detectable in high precision Kepler photometry (e.g. Kallinger et al., 2014).

Faculae, and their chromospheric counterparts, plages (see Fig. 3.4), are brighter than the average stellar surface, but their temperature contrast is less pronounced than for spots (Schrijver, 2002; Thomas and Weiss, 2012). Therefore, even though plages occupy a larger area of the stellar surface, they induce a small flux effect compared to that of a spot (Meunier et al., 2010a). In the Sun, and likely also in low-activity stars (Shapiro et al., 2016), spots are always surrounded by plages, but the converse is not true: plages can exist on their own (Thomas and Weiss, 2012).

Besides the flux effect, and due to their strong local magnetic fields, active regions also inhibit convective motions just below the stellar surface (Title et al., 1987; Hanslmeier et al., 1991). This suppresses part of the convective blueshift originating from granulation (Gray, 2009). Both spots and faculae therefore appear redshifted in comparison to the quiet photosphere. As they appear and disappear from the visible stellar hemisphere, the *suppression of convective blueshift* can induce a net RV variation of several $\mathrm{m{\cdot}s^{-1}}$ (Meunier et al., 2010a; Meunier and Lagrange, 2013; Dumusque et al., 2014a), with plages showing a much more pronounced effect, qualitatively different from that of spots (see Meunier et al., 2010a; Dumusque et al., 2014a).

On a shorter timescale (between 5 and 15 minutes), the oscillations caused by the propagation of pressure waves on the stellar interiors (Leighton et al., 1962; Bedding et al., 2001; Kjeldsen et al., 2005), can also induce RV variations. In amplitude, these variations tipically vary between $10\,\mathrm{cm{\cdot}s^{-1}}$ and about $4\,\mathrm{m{\cdot}s^{-1}}$ (e.g. Bouchy et al., 2005), depending on the stellar type (Christensen-Dalsgaard, 2004).

One other important source of RV variation induced by stellar activity are long-term stellar magnetic cycles. The Sun undergoes a cyclic regeneration of its magnetic field, with a reversal in polarization every 11 years, and an overall periodicity of 22 years. This cycle is

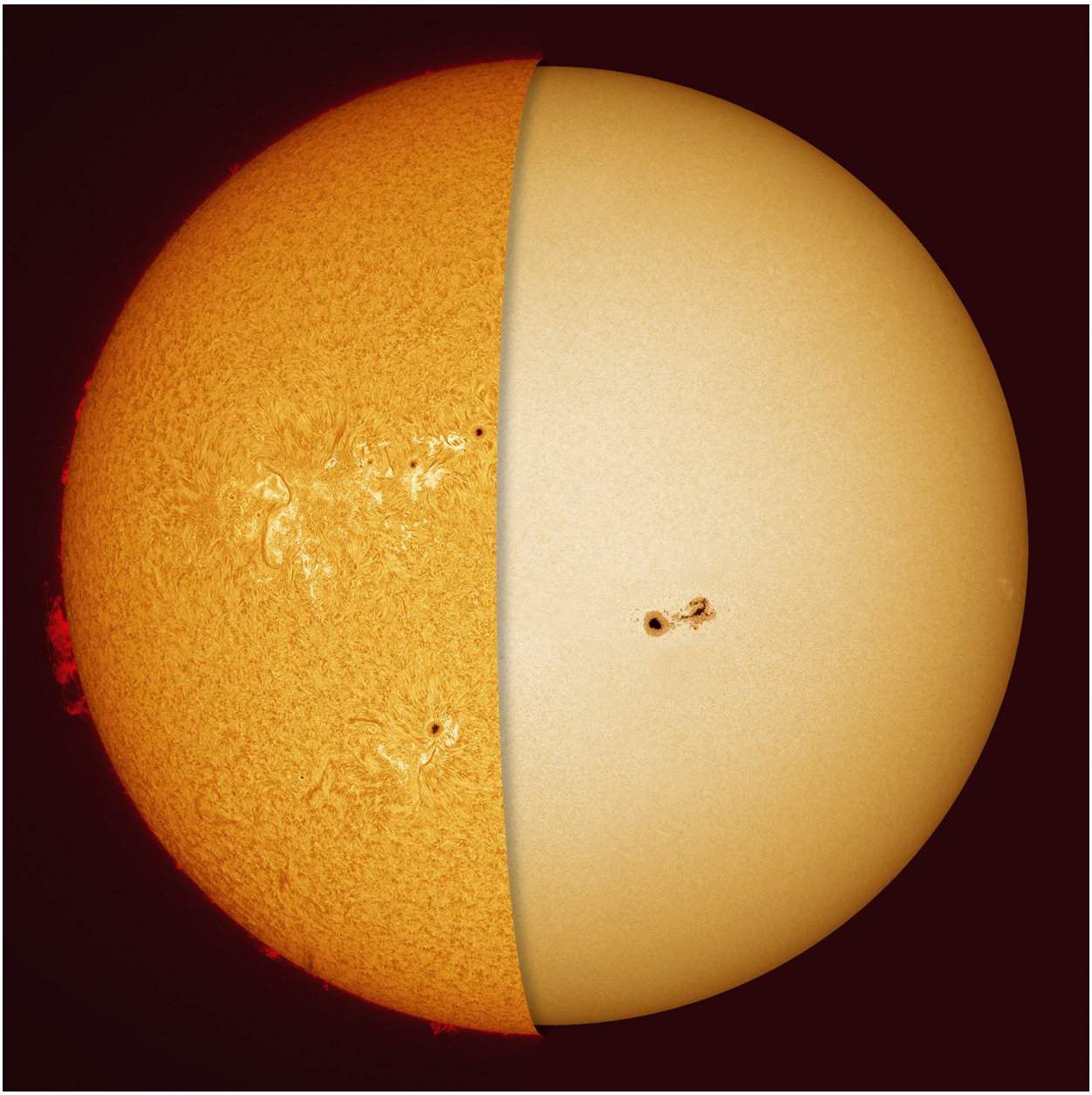


Figure 3.4: A stunning composite picture of the Sun. On the right, the view in white light, showing the solar photosphere. Here we see sunspots and brighter faculae at the limb. On the left is the Sun at the wavelength of Hα, representing the chromosphere. In this wavelength we see prominences, filaments, granulation, sunspots and plages. The artistic effect depicts how the chromosphere is located above the photosphere. Created by Alan Friedman, see avertedimagination.com.

characterized by an increase and decrease in the number and surface area of sunspots, together with other solar activity indicators which also vary periodically (Charbonneau, 2010; Hathaway, 2015).

Many solar-type stars show the same magnetic activity cycles (Baliunas et al., 1995; Oláh et al., 2009), and these cycles have timescales comparable to the orbital periods of Jupiter-like planets. Their impact on precise RVs was studied by Santos et al. (2010a) and Lovis et al. (2011a). The second work in particular found RV variations of up to $25\,\mathrm{m\cdot s^{-1}}$, but which could be diagnosed (and likely corrected) using simultaneous Ca II H&K measurements, for example. Several examples have since been analysed in detail (see for example Dumusque et al., 2011a; Carolo et al., 2014; Díaz et al., 2016; Johnson et al., 2016).

3.2 CORRECTING FOR STELLAR ACTIVITY

Because stellar activity can generate RV variations that prevent the detection and accurate measurement of planetary signals, active stars are often discarded from RV surveys (e.g. Fischer et al., 2016, and references therein). But in a few cases, we have been misled towards the spurious detection of planets around stars with low or moderate levels of activity (e.g. Queloz et al., 2001; Desidera et al., 2004; Figueira et al., 2010; Carolo et al., 2014; Kane et al., 2016, just to name a few).

Moreover, the discovery of small transiting planets around active stars, for example with the CoRoT and Kepler missions (e.g. Barros et al., 2014; Barclay et al., 2015), meant that one had to deal with the stellar activity contaminations. A number of methods were developed to do so, either based on a custom observational strategy (Dumusque et al., 2011b,c), auxiliary spectral indicators of activity (Queloz et al., 2001; Bonfils et al., 2007; Desort et al., 2007), simultaneous photometric observations (Boisse et al., 2009; Haywood et al., 2014; Giguere et al., 2016) or careful analysis of the RV timeseries (e.g. Hatzes et al., 2010; Hatzes, 2014). In this Section, I present a few of these methods.

3.2.1 *Observational strategies*

Dumusque et al. (2011b,c) investigated different observational strategies and calendars, with the objective of reducing the sources of RV noise coming from stellar activity. They considered both real and synthetic RV measurements, and tested which observational strategy provides the smallest RV variation due to stellar noise, taking into account the contributions of oscillations, granulation, active regions, and instrumental errors.

The observational strategy used previously in HARPS-GTO observations – one measurement of 15 min per night – already succeeded in reducing stellar oscillation noise, but was inefficient to average out

granulation phenomena, due to the different timescales involved. By increasing the number of measurements per night and binning those observations (thereby sampling the granulation timescales), one reduces the granulation noise more efficiently, improving the planetary detection limits (Dumusque et al., 2011c).

When further taking into account the contribution of stellar spots, the most efficient strategy to average out all kinds of noise consists in taking one RV measurement (10 min) three times per night, with a spacing of two hours, every third night, 10 days every month (Dumusque et al., 2011b).

It is clear that these observational strategies require not only a big investment in telescope time, but also careful planning and scheduling of the observations. This means that only a handful of stars ever get close to the optimal sampling strategy. Nevertheless, exposures of at least 15 minutes are now the norm when observing with spectrographs such as HARPS or HARPS-N.

3.2.2 *Line-profile indicators*

As we saw before, the CCF contains information not just about the RV of the star, but also about stellar activity. In order to use this information to correct the RVs, a number of *line-profile indicators* have been proposed. These indicators are usually moments of the CCF, or measures of its asymmetry, which show correlations with the RVs for signals created by activity.

Queloz et al. (2001) first proposed the bisector inverse slope (BIS). The bisector of the CCF is the line which divides it in two halves of equal equivalent width. The effect from the convective blueshift produced by granulation, is to bend the bisector towards the top (e.g. Gray, 1989, and references therein). The presence of active regions on the stellar surface will further distort this shape (see Fig. 3.3).

The bisector has been successfully used to identify or dismiss stellar activity signals (Santos et al., 2002; Melo et al., 2007; Boisse et al., 2009), but it has a number of limitations (see e.g. Figueira et al., 2013). For example, Prato et al. (2008) studied a group of young stars with RV variations that likely resulted from spots, and did not find significant correlations with the bisector span (see also Carolo et al., 2014). In addition, if the rotational velocity of the star is lower than the resolution of the spectrograph, variations in RV and bisector spans will not correlate (Santos et al., 2003; Desort et al., 2007).

For stable spectrographs, such as HARPS, the full width at half maximum (FWHM) and the contrast of the CCF can also reveal profile variations caused by stellar activity (Boisse et al., 2011). The FWHM, in particular, sometimes shows a better correlation with the activity-induced signals seen in photometry than the BIS, for slowly rotating

In non-stabilized spectrographs, the FWHM will depend essentially on the instrumental profile, which is not stable as function of time.

stars (e.g. Queloz et al., 2009; Santos et al., 2014).

In addition to these, a number of other indicators were proposed:

- The $V_{\rm span}$ from Boisse et al. (2011), consists in fitting a Gaussian profile to the top and bottom parts of the CCF. The indicator is then defined as the difference between the two center velocities.
- BiGauss, proposed by Nardetto et al. (2006) and further tested in Figueira et al. (2013), which fits an asymmetric Gaussian profile to the CCF. The diagnostic is defined as the difference between the RV found with the symmetric and asymmetric Gaussian profiles.
- The $V_{\rm asy}$ indicator (Figueira et al., 2013) consists in computing the difference in spectral information contained in the blue and red wings of the CCF.[4]
- $W_{\rm span}$: (Santerne et al., 2015) fits a Gaussian profile to the red and blue wings of the line profile. The asymmetry indicator is defined as the difference between the two center velocities.

[4] After a correction, Figueira et al. (2015), advise against the use of $V_{\rm asy}$.

All these indicators[5] can be obtained from the stellar spectrum without additional observational effort. However, the response of these indicators to the active region distribution on the stellar surface is complex, as is their relation with the activity-induced RV signals. This means that a single indicator is often not enough to flag a stellar activity signal, except when a very clear and significant correlation with the RVs is present.

[5] For a visual representation of the line profile indicators see Santerne et al. (2015), their figure A.1.

The usual strategy when using these indicators is therefore to compute all of them every time we suspect of activity-induced RV variations. The FWHM, contrast, and BIS span are automatically computed by the HARPS pipeline, and are readily available. Figueira et al. (2013) provided a software suite to compute other indicators such as the $V_{\rm span}$, BIS+ and BIS- (alternative formulations of the BIS). Finally, the (computationally simple) calculation of the $W_{\rm span}$ is described in some detail in Santerne et al. (2015).

A typical analysis consists in calculating the Lomb-Scargle periodograms of the activity indicators and comparing any significant peaks with those found in the RVs themselves. Sometimes, the periodograms can also be used to identify the stellar rotation period, although care must be taken not to falsely detect the rotation period harmonics (e.g. Boisse et al., 2011).

A more intricate analysis of some activity indicators was recently presented by Rajpaul et al. (2015), who used Gaussian process regression to model simultaneously the joint variations of the RVs, the FWHM, BIS span, and $\log R'_{\rm HK}$. This model rests on strong assumptions on how the different indicators are related to the spot distribution. Nevertheless, it allowed for the activity component to be constrained and disentangled from planetary signals.

In the case of M dwarfs, activity indicators derived from the CCF have also proved useful in identifying signals caused by stellar activity (Robertson and Mahadevan, 2014). However, a few recent debates in the literature have highlighted our lack of a complete understanding about how the indicators relate to activity (Anglada-Escudé and Tuomi, 2015; Robertson et al., 2015; Anglada-Escudé et al., 2016b).

3.2.3 *Indicators of chromospheric activity*

The temperature difference between stellar plages and the surrounding regions changes the rate of atomic transitions within these active regions. Because of this effect, the relative depth of the Hα line and the emission in the cores of Ca II H&K lines and Na I D lines, have been found to be good indicators of magnetic stellar activity (e.g. Mallik, 1997; Saar et al., 1998; Cincunegui et al., 2007; Gomes da Silva et al., 2011; Robertson et al., 2013).

Vaughan et al. (1978) first used the dimensionless S-index, as an indicator of stellar activity. The S-index is calculated as the ratio of the emission in the line cores of Ca II H&K, to that in two nearby continuum bandpasses on either side of the lines (Vaughan et al., 1978; Santos et al., 2000; Hall, 2008):

$$S = \frac{H + K}{C_{3900} + C_{4000}}, \tag{3.1}$$

where H and K are the total fluxes at the cores of the lines, and C_{3900} and C_{4000} the total fluxes in the continuum regions centred on 3900Å and 4000Å. The S index provides a measure of activity for the Sun or a given star, but it contains a color term (coming from the continuum reference bandpasses) and also a photospheric contribution in the line core bandpasses. Therefore, the S-index cannot be used to compare the chromospheric activity of stars with different colours.

To overcome these limitations, Noyes et al. (1984a) proposed the $R'_{\rm HK}$ index, defined as the fraction of a star's bolometric luminosity radiated as chromospheric H and K emission (Hall, 2008):

$$\log R'_{\rm HK} = 1.34 \times 10^{-4} \cdot C_{ef} \cdot S \tag{3.2}$$

with

$$C_{ef} = 1.13(B-V)^3 - 3.91(B-V)^2 + 2.84(B-V) - 0.47\,,$$

with $B - V$ the color excess of the star (see also Dumusque, 2016). In this way, the $R'_{\rm HK}$ is corrected for the stellar photospheric continuum emission using a callibration that depends on the spectral type of the star (Rutten, 1984; Martínez-Arnáiz et al., 2010), and can be used to compare stars of different spectral types (see also Lovis et al., 2011a). The logarithmic version of the index, denoted $\log R'_{\rm HK}$, is very often

used as a stellar activity proxy in RV analysis. Its numerical value ranges tipically from about -4.4 to -5.1 for very active to inactive stars (e.g. Hall, 2008). For the Sun, the $\log R'_{\rm HK}$ takes the value -4.895 (Lockwood et al., 2007; Shapiro et al., 2013). A division between *active* and *inactive* stars is ususally considered at $\log R'_{\rm HK} = -4.75$, corresponding to the limit of the so-called Vaughan-Preston gap (see Vaughan and Preston, 1980; Hall, 2008).

During a solar-like magnetic cycle, as the number of spots on the stellar surface increases, the inhibition of the convective blueshift makes the stellar spectrum appear redshifted, causing a positive radial velocity. This means that, in the presence of such a magnetic cycle, we expect a positive correlation between the $\log R'_{\rm HK}$ and the long-term RV variations.

Therefore, as in the case of the CCF indicators, we typically look for common periodicities or correlations between the timeseries of RVs and $\log R'_{\rm HK}$ (e.g. Bonfils et al., 2007; Johnson et al., 2016). In more sophisticated models, the $\log R'_{\rm HK}$ variation can also be used to set a white noise term in the RVs, as in Díaz et al. (2016). Other activity proxies can also be considered, based on the Hα, Na I D or He I D3 lines, for example (Santos et al., 2010a).

3.3 MODELLING ACTIVITY-INDUCED RADIAL-VELOCITY VARIATIONS

We have seen how stellar activity can make planet detection harder, and how different techniques can be used to correct for it. In essence, the RV signals induced by stellar activity can be seen as very complex noise patterns. Therefore, another way to disentangle activity from planets is to consider more sophisticated noise models, that aim to represent all the physical components at play. Building upon the model introduced in Chapter 2, I have developed such a framework for planet detection in active stars.

In this section, I start by introducing Gaussian process regression, the mathematical tool which will be used to model the quasi-periodic activity signals. An in-depth analysis of the HARPS data for CoRoT-7 is then presented, which was first published as Faria et al. (2016a). Some other examples of the application of this model to stars from the metal-poor Large Program will be shown in Chapter 4.

3.3.1 *Gaussian processes*

When reading about Gaussian processes, not a single reference will fail to cite Rasmussen and Williams (2006) and probably mention their definition:

> A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.
>
> (Rasmussen and Williams, 2006)

As the rest of their book, this definition is "primarily intended for graduate students and researchers in machine learning at departments of Computer Science, Statistics and Applied Mathematics." Therefore, it is to be expected that the definition does not convey a great deal of insight to an astronomer, at least on a first read. In that regard (of insight), David MacKay did a wonderful job in MacKay (1998, 2006). There he explains how the simple Gaussian distribution can be generalised and used for nonlinear regression. Here I follow closely his explanation.

Let us start with the 1-dimensional Gaussian distribution (the normal distribution, or the bell curve). If the variable X_1 follows a Gaussian distribution, its probability density function (pdf) is

$$p(X_1 \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{3.3}$$

with μ and σ the mean and standard deviation of the distribution, respectively. Figure 3.5 shows two such probability densities, for $(\mu = 0, \sigma = 1)$, and for $(\mu = -3, \sigma = 0.6)$, and 50 random samples from each distribution. Note that a random sample in this case is simply one number.

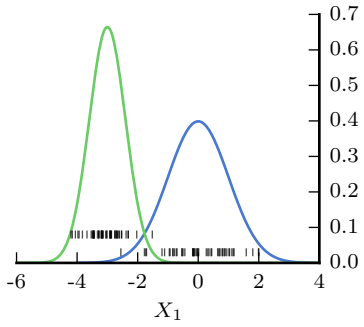


Figure 3.5: Two examples of the 1-dimensional Gaussian distribution and a representation of 50 random samples from the two distributions.

Now consider the 2-dimensional case. If variables X_1 and X_2 follow a *bivariate* Gaussian distribution, their (joint) pdf is

$$p(\mathbf{X} \mid \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{k/2}\sqrt{\det \Sigma}} \exp\left[-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X}-\boldsymbol{\mu})\right] \tag{3.4}$$

where $\mathbf{X} = [X_1, X_2]^T$, k is the dimension of $\mathbf{X}$, in this case 2, $\boldsymbol{\mu}$ is the vector of the means, $\boldsymbol{\mu} = [\mu_1, \mu_2]$, and Σ is the covariance matrix, a 2×2 matrix with entries

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

The variable X_1 follows a 1-D Gaussian distribution, with mean μ_1 and standard deviation σ_1, and similarly for X_2. The *covariance* between the two variables is σ_{12} and their *correlation* is given by $\sigma_{12}/(\sigma_1 \cdot \sigma_2)$. Fig. 3.6 shows the bivariate probability density, for $\mu_1 = 0, \mu_2 = 4$, $\sigma_1 = \sigma_2 = 1$, and $\sigma_{12} = 0$. The figure also shows 500 random samples from the distribution. These 500 *samples* now mean 500 sets of values (x_1, x_2).

Now consider the 3-dimensional case[6] for variables X_1, X_2 and X_3. The joint pdf is the same as in Eq. (3.4) with $\mathbf{X} = [X_1, X_2, X_3]^T$, $k = 3$, $\mu = (\mu_1, \mu_2, \mu_3)$ and Σ a 3×3 matrix. This 3-dimensional distribution is harder to represent visually but, other than that, is a

[6] This is going somewhere.

simple generalisation of the 1-D case. Coincidentally, if we do try to represent it in a figure, we might understand Gaussian processes!

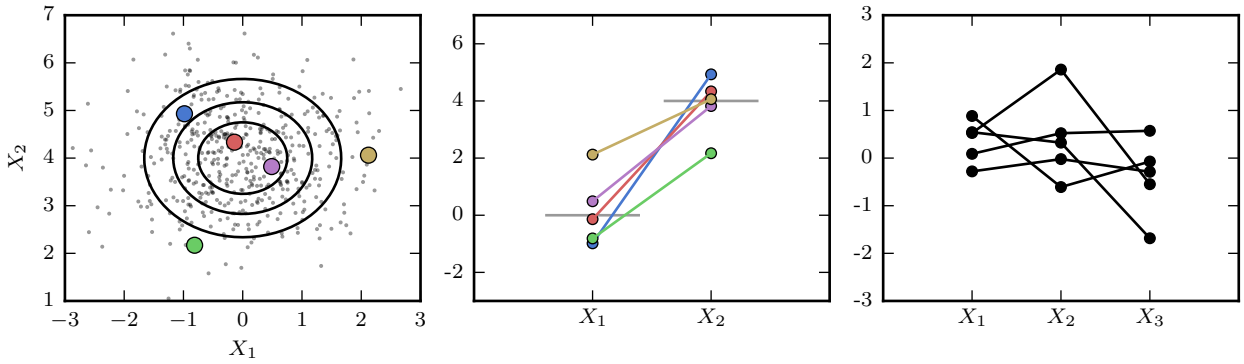


Figure 3.6: *Left*: a 2-dimensional Gaussian distribution, represented with contour levels and 500 random samples. The extra 5 coloured samples are also shown in the middle panel. *Middle*: new representation of the Gaussian distribution, with variables X_1 and X_2 in the x-axis and their values in the y-axis. The two horizontal lines represent each variable's mean value. *Right*: the same representation, now of a 3-dimensional Gaussian distribution. Note that in the middle panel the apparent "correlation" comes only from the difference in the mean value for each variable.

We can use the trick that is laid out in Fig. 3.6, middle panel, for the bivariate distribution. Each of the two variables is represented in the abscissae, and a value that each variable takes is in the ordinate axis. Note that, in this representation, a *sample* from the bivariate Gaussian distribution is a *line* connecting two values of the variables (x_1, x_2). The figure shows 5 such samples, color-coded to correspond to those same samples in the left panel.

The new representation allows us to visualise the 3-D distribution as well, as shown in the right panel of Fig. 3.6. Three variables are now in the x-axis, and lines connecting their values represent samples from the distribution. The panel represents 5 samples drawn from a 3-D distribution with all means set to 0, unit standard deviations, and zero covariances.

This new way to visualise a multivariate Gaussian distribution works for any number of dimensions, or variables. For example, we can set $k = 50$ and represent a 50-dimensional Gaussian distribution. Three samples from three different distributions are shown in Fig. 3.7. The difference between the three panels is the covariance matrix (all means are again set to zero). In the top panel, Σ is the $k \times k$ identity matrix, in the middle panel the value of the correlation decreases with the variable index number, and in the bottom panel the correlations are periodic.

All representations so far show one or more samples from each distribution. But we can also represent the distributions by their mean and 95% (iso-)probability intervals, corresponding to two standard deviations. We now consider a 500-dimensional Gaussian distribution and plot its mean and 3 random samples in the left panel of Fig. 3.8. The shaded region denotes twice the standard deviation at each input value $X_1, \ldots, X_{500}$.

Let us imagine that this distribution is taken as a prior for the 500 variables. What would happen if we observed some data, say two of

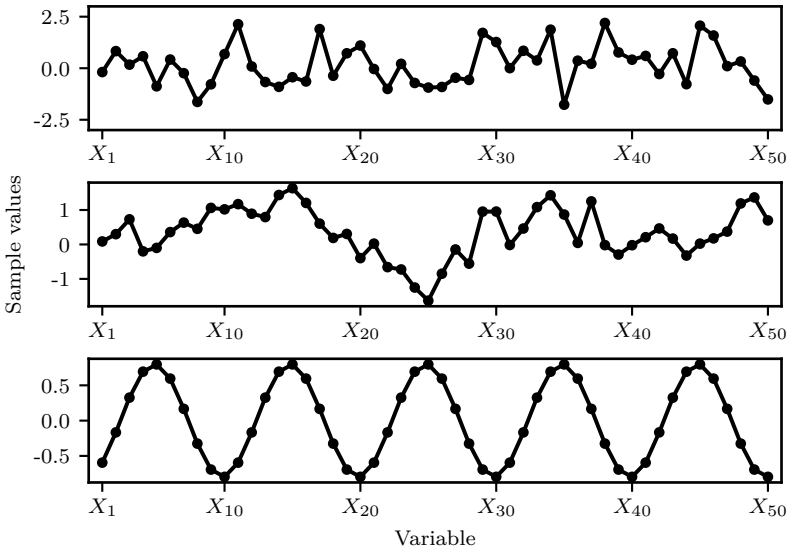


Figure 3.7: Samples from a Gaussian distribution with $k = 50$ and different covariance matrices.

those variables? In the right panel of Fig. 3.8, the variables X_{70} and X_{160} have been *observed* to have the values 0.9 and -1.6, respectively. The dashed lines in the figure show samples which are consistent with the data, and the solid line depicts the mean value of those samples. The shaded region still denotes twice the standard deviation at each X. Notice that the uncertainty (the variance of the samples) is reduced close to the observations.

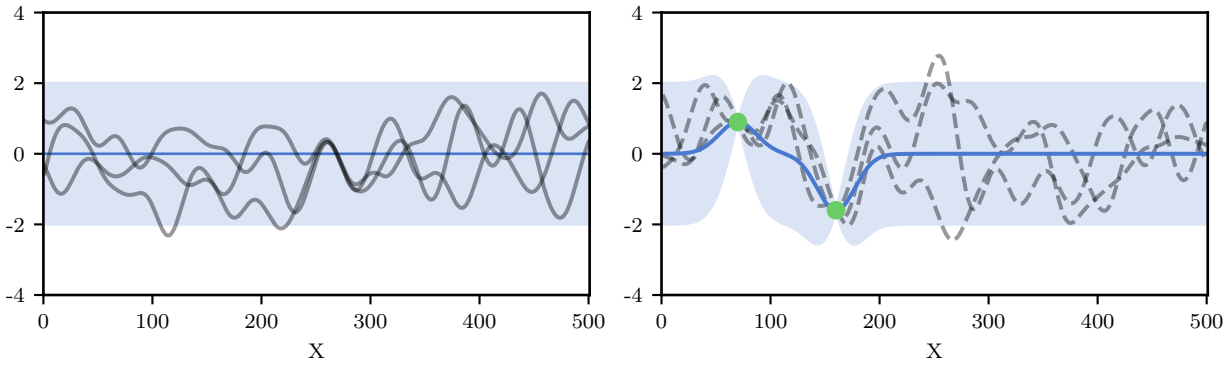


Figure 3.8: Illustration of GP regression. The left panel shows three random draws from a 500-dimensional Gaussian distribution. The blue solid line is the mean value of the distribution and the shaded region represents two times the standard deviation for each variable X. The right panel depicts how the distribution changes when variables X_{70} and X_{160} are observed.

The reader will note that we can consider as many different variables as we want. In fact, we can even make the leap from random variables to functions. Without any mathematical rigour, a function is just a very long (infinite) vector, with each entry giving the function value $f(x)$ at a particular input x. But how to deal with an infinite number of variables? It turns out that when we ask only for the properties of the function at a finite number of points, the answer is the same whether we ignore the infinitely many other points, or we take them all into account.

We can now see Fig. 3.8 in a new light. The x-axis does not have to correspond to a fixed number of random variables, but can instead correspond to an input space x. A Gaussian process is written as

$$f(x) \sim \mathcal{GP}(m(x), \mathcal{K}(x, x')) \tag{3.5}$$

and corresponds to a distribution over functions $f(x)$, where $m(x)$ is the *mean function* and $\mathcal{K}(x, x')$ is the *covariance function*, sometimes also called a *kernel*.

The Gaussian process can be a *prior* distribution over functions which, when combined with data, provides a *posterior* distribution over functions. We are using the Gaussian distribution to perform Bayesian nonlinear regression!

The properties of the GP are set by its mean and covariance functions. So far I mostly considered $m(x) = 0$ in the examples, but nothing restricts the mean function from being anything else. In fact, any function of x is acceptable, even if parameterized by some set of parameters θ. We could consider for example a linear function $m(x) = a\,x + b$. The covariance function also helps in further specifying what kind of functions are included in the GP distribution. $\mathcal{K}(x, x')$ can set many characteristics of the functions, such as their smoothness and stationarity (whether they look similar at all x locations). How fast the functions vary – that is their characteristic length-scales – can be defined by both the mean and the covariance.

There are many covariance functions available, the only restriction being that the kernel $\mathcal{K}(x, x')$ has to produce a positive semidefinite covariance matrix (Rasmussen and Williams, 2006). I will only consider stationary covariance functions, that is, functions of $x - x'$. These functions are invariant to translations in the input space.

For the examples of Figures 3.7 and 3.8, I have used the following forms for the covariance kernels:

- the diagonal kernel (δ is the Kronecker delta)

$$\mathcal{K}(x, x') = \eta_1\, \delta_{x,x'} \tag{3.6}$$

- the exponential kernel

$$\mathcal{K}(x, x') = \eta_1 \exp\left(-\frac{x - x'}{\eta_2}\right) \tag{3.7}$$

- the cosine kernel

$$\mathcal{K}(x, x') = \eta_1 \cos\left(\frac{2\pi}{\eta_2}|x - x'|\right) \tag{3.8}$$

- the squared exponential kernel

$$\mathcal{K}(x, x') = \eta_1 \exp\left[-\frac{(x-x')^2}{2\eta_2^2}\right] \tag{3.9}$$

Each covariance function has a given number of parameters that define its properties, such as amplitudes, length-scales or periods. Sums and products of covariance functions are also valid covariance functions. The covariance matrix is built by evaluating the covariance function at each of the input points:

$$\Sigma_{ij} = \mathcal{K}(x_i, x_j) \tag{3.10}$$

3.3.2 *Quasi-periodic stellar activity signals*

Now that I have introduced GPs, we can make the link to stellar activity. It is clear that activity signals cannot be easily described analytically: a complete description would require a complete and detailed knowledge of the active region distribution and evolution, the temperature contrast of every active region, and stellar parameters such as inclination and limb darkening. But we can make some assumptions about the form of such signals and use GPs to model them.

Besides assuming the signals will be continuous and smooth, stellar rotation will induce a periodicity or quasi-periodicity, as the active regions evolve and cycle in and out of view on the stellar surface. We can make the hypothesis that a *quasi-periodic* kernel could model these variations as a function of time t:

$$\mathcal{K}_{\rm QP}(t_i, t_j) = \eta_1^2 \exp\left[-\frac{(t_i-t_j)^2}{2\eta_2^2} - \frac{2\sin^2\left(\frac{\pi(t_i-t_j)}{\eta_3}\right)}{\eta_4^2}\right]. \tag{3.11}$$

Note that we are not modeling the RV variations directly, but instead the correlation of consecutive RV observations. This is a subtle but key point.

This kernel represents some of our expectations for the activity-induced RV signals (Haywood et al., 2014; Grunblatt et al., 2015). First, the magnitude of the correlations decays on a timescale of η_2 days (first term inside the exponential) and has a periodic variation, with period η_3 days (from the second term in the exponential). These properties relate to the evolution timescale of active regions and to the rotation period of the star. The ratio η_2/η_3, and to a certain extent η_4 (which is adimensional), control the relative importance of the periodic and decaying components. The amount of high-frequency structure in the periodic component is set by η_4. The parameter η_1, measured in $\mathrm{m \cdot s^{-1}}$, represents the amplitude of the correlations, with η_1^2 being the variance of the GP functions.

In Fig. 3.9, I show SOAP 2 simulations (Dumusque et al., 2014a)[7] of a solar-type star with 5 spots positioned at random locations in

[7] SOAP (and its second version, SOAP 2) is a simulation code that estimates the photometric and RV variations induced by active regions (see Boisse et al., 2012; Oshagh et al., 2013; Dumusque et al., 2014a).

the stellar surface. The plot shows the total RV effect over 3 rotation periods. Below, a sample from the conditional distribution of a GP with a quasi-periodic kernel mimics almost perfectly the signal, while other kernels fail to do so.

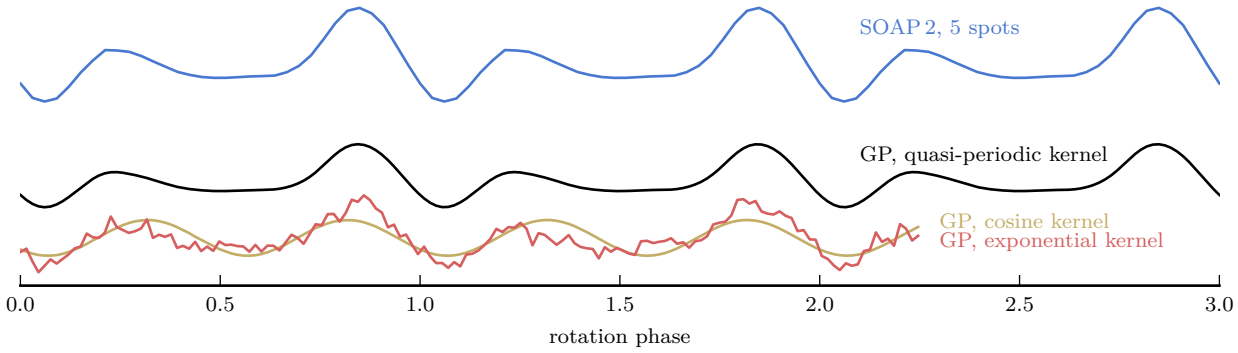


Figure 3.9: Simulation of the RV effect of 5 spots over three rotation periods on a solar-type star. A sample from the conditional distribution of a GP with a quasi-periodic kernel mimics the signal but the cosine or the exponential kernel present too large or too short variations.

When searching for planets in the RVs of active stars, we can consider stellar activity to be the noise affecting the interesting signal coming from the planets. Therefore, we can use a GP with the covariance function of Eq. (3.11) and a mean function given by a sum of N_p Keplerians, Eq. (2.64), to model the RVs.

The additive white noise component s, that we considered back in Section 2.3.5, can be taken into account with a diagonal kernel. Then, the full specification of the model is

$$m(t) = \sum_{\text{all planets}} v_{r,\,\text{model}}\,(t) \tag{3.12}$$

$$\mathcal{K}(t_i, t_j) = \mathcal{K}_{\text{QP}}(t_i, t_j) + s^2\delta_{ij} \tag{3.13}$$

$$v_{r,\,\text{obs}}\,(t) \sim \mathcal{GP}(m(t), \mathcal{K}(t_i, t_j)) \tag{3.14}$$

This model is an extension of the one presented in Section 2.3.4, and can be represented by the probabilistic graphical model of Fig. 3.10, including the four additional parameters of the covariance function. There will be only two differences in estimating this model, compared to the original one without the GP component: the calculation of the likelihood and the extra priors for η_1, η_2, η_3 and η_4.

The log-likelihood is calculated with the expression

$$\ln p(D \mid \Theta, \mathcal{I}) = -\frac{1}{2}\mathbf{r}^T\,\Sigma^{-1}\,\mathbf{r} - \frac{1}{2}\ln\det\Sigma - \frac{N}{2}\ln 2\pi, \tag{3.15}$$

where $\mathbf{r}$ is the residual vector given by

$$\mathbf{r} = v_{r,\,\text{obs}} - v_{r,\,\text{model}} \tag{3.16}$$

calculated for all data points.

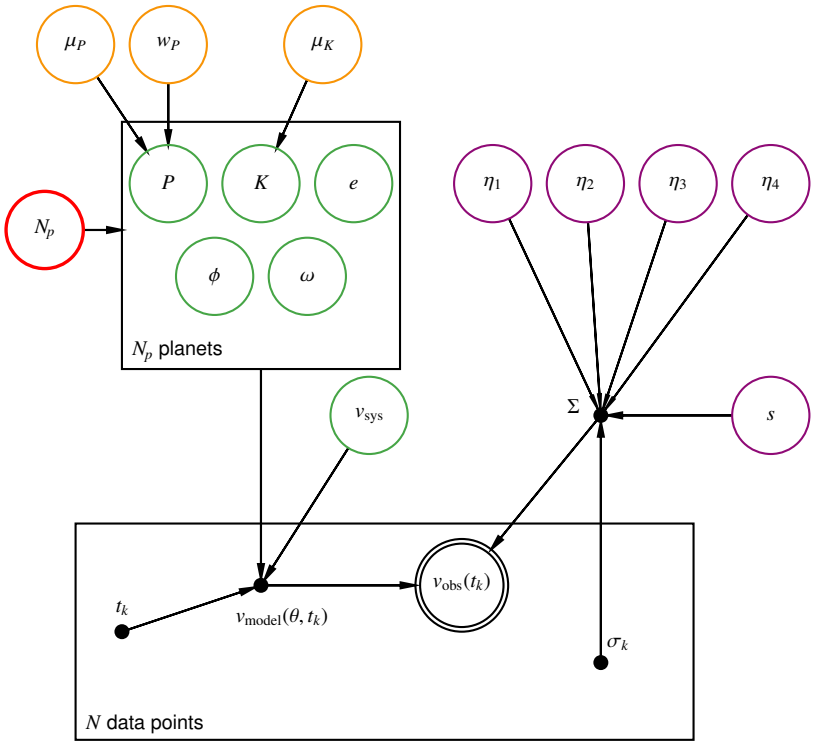


Figure 3.10: Full graphical model when including stellar activity. The covariance matrix Σ is a deterministic function of the hyperparameters η_1, η_2, η_3 and η_4, the jitter s, and the observed uncertainties σ_k. Compare with Fig. 2.12. Adapted from Faria et al. (2016a).

The typical priors we use for the GP parameters are shown in Table 3.1. Note that some of these parameters are physically interpretable. The parameter η_3 corresponds to the stellar rotation period and η_2 to a typical timescale for evolution of the active regions. Therefore, the priors for these parameters can be set individually for a given star, using other available information.

As a typical example, if the rotation period is known, a more restrictive prior for η_3 (such as a Gaussian distribution) could be considered instead of the reference prior. The stellar rotation period can sometimes be estimated from photometric observations (e.g. McQuillan et al., 2014; Aigrain et al., 2015) or from activity-rotation relations (Noyes et al., 1984b; Mamajek and Hillenbrand, 2008).

GP parameters			
η_1	amplitude of covariance	$\mathcal{LU}(0.1, 50)$	$\mathrm{m \cdot s^{-1}}$
η_2	aperiodic timescale	$\mathcal{LU}(1, 100)$	days
η_3	correlation period	$\mathcal{U}(10, 40)$	days
η_4	periodic scale	$\mathcal{LU}(0.1, 10)$	

Table 3.1: Meaning and reference prior distribution for parameters of the quasi-periodic kernel in the GP model. Note that the prior for η_3 is only a default one and its limits can be set for each star, based on previous information. Symbols have the same meaning as in Table 2.1.

The typical lifetimes of sunspots and starspots are of about a month for solar-type stars (but longer for M-dwarfs and giant stars; Hussain, 2002; Schrijver, 2002), even if a wide range of values has been observed, ranging from a few days for the shortest lived active regions to years for more active stars (Bradshaw and Hartigan, 2014).

For η_1 and η_4, it is harder to come up with physically-motivated

priors, and we use general log-uniform priors for these parameters, since we are a priori unsure of their order of magnitude. Nevertheless we expect η_1 to represent the magnitude of the stellar activity signals present in the RVs, and η_4 to be of the order of 1. Some RV curves may require $\eta_4 \sim 0.5$, allowing for structure to develop at half the rotation period.

* ◯ *

At this point, it is important to clarify the steps involved in the estimation of the parameters of our model. As before, we use the birth-death MCMC of Brewer (2014), integrated within the DNS algorithm, so it suffices to describe the calculations performed at each step in the MCMC. Consider a set of parameters Θ at a given step:

1. using the subset of planet parameters from Θ, calculate $v_{r,\,\mathrm{model}}$ and $\mathbf{r}$, at the observed times.
2. using the subset of GP parameters from Θ, build the covariance matrix Σ using Eq. (3.10) and the quasi-periodic kernel, Eq. (3.11). Add s^2 to the diagonal entries of Σ.
3. invert Σ and calculate its determinant.
4. calculate the log-likelihood, Eq. (3.15).
5. calculate the prior pdf for the current value of Θ.

Note that the covariance matrix Σ is $N \times N$, and so step 3 above can be computationally demanding. In general, the inversion of the matrix is done using its Cholesky decomposition (because the matrix is positive semidefinite), which requires about N^3 operations. For somewhat large values of N this becomes a severe limitation, but for typical RV datasets, with up to a few hundreds of points, the calculation is still feasible. We note, however, that the development of faster algorithms for GP analysis is a very active research area (Ambikasaran et al., 2014; Foreman-Mackey et al., 2017), with very promising results.

Originally published by Cholesky (1910), the method is discussed in Brezinski (2005) and presented in any standard linear algebra textbook.

As before, we will obtain samples from the posterior distributions for N_p, for the planet parameters, and now the GP parameters, simultaneously. The posterior for η_3 will represent a constraint on the stellar rotation period, which is obtained only using RV measurements.

3.4 APPLICATION TO CoRoT-7

I now apply the method described in the previous section to HARPS observations of the planet-host star CoRoT-7 .

CoRoT-7 b was first announced by Léger et al. (2009) and was the first super-Earth to have a measured radius ($R = 1.68 \pm 0.09 R_\oplus$). Its

orbital period was estimated from the transits in the CoRoT lightcurve to be $P_b = 0.85359 \pm 0.00003$ days (Léger et al., 2009).

A second non-transiting planet, with an estimated orbital period of $P_c = 3.698 \pm 0.003$ days, was detected in a follow-up RV campaign (Queloz et al., 2009) and a more disputed detection of a third planetary signal was reported by Hatzes et al. (2010) and Tuomi et al. (2014a).

CoRoT-7 's lightcurve (in 2008-2009) shows out-of-transit modulations of up to 2%. This can be compared to the Sun, in which the greatest recorded variations in irradiance are of 0.34% (Kopp and Lean, 2011). In the case of CoRoT-7 this photometric variation translates into a high RV variation, even if this is not always the case.

Due to the high activity level, this system has generated a wealth of discussion, resulting in different estimates for the masses of the two planets and uncertainty about the presence of the third one (Hatzes et al., 2010; Lanza et al., 2010; Boisse et al., 2011; Ferraz-Mello et al., 2011; Hatzes et al., 2011; Pont et al., 2011). A review of these works would be too lengthy and, instead, Table 3.2 shows a compendium of the most relevant results.[8] In Fig. 3.11, the mass estimates from these works are compared with the ones derived using our model (see Table 3.3 below).

[8] See also Hatzes et al. (2011), for different mass determinations of CoRoT-7 b.

	CoRoT-7 b		CoRoT-7 c		CoRoT-7 d	
	mass [$M_\oplus$]	period [days]	mass [$M_\oplus$]	period [days]	mass [$M_\oplus$]	period [days]
Boisse et al. (2011)	5.7 ± 2.5	‡	13.2 ± 4.1	3.695 ± 0.02	nd	
Ferraz-Mello et al. (2011)	8 ± 1.2	‡	13.6 ± 1.4	‡	na	
Pont et al. (2011)	2.3 ± 1.8	0.854	nd		nd	
Hatzes et al. (2010)	6.9 ± 1.4	0.853 59(59)	12.4 ± 0.42	3.691 ± 0.0036	16.7 ± 0.42	9.021 ± 0.019
Hatzes et al. (2011)	7.42 ± 1.21	‡	na		na	
Haywood et al. (2014)	4.73 ± 0.95	0.853 591 65(5)	13.56 ± 1.08	3.70 ± 0.02	nd	

Table 3.2: Estimates for the masses and orbital periods of CoRoT-7 's planetary companions, derived from different analysis. Notes: *nd* means the planet was not detected and *na* means the planet is not analysed in detail. Parameters marked with ‡ were fixed to previously-determined values. The numbers in brackets represent the uncertainty in the last digits of the value.

It is important to note that these different analyses of the CoRoT-7 RVs rely not only on different methods, but often on additional data, such as spectroscopic activity indicators or photometry. In particular, the orbital period of CoRoT-7 b was not always estimated from the data, but instead fixed to the transit-derived value, which is assumed to be much more precise than any value that can be derived from RV observations.

In 2012, CoRoT and HARPS were used simultaneously to observe CoRoT-7 and to help with the mass determinations (Barros et al. 2014; Haywood et al. 2014, hereafter H14). These simultaneous observations were analysed by H14 with a model that is similar to the one I present here but which considers information from the CoRoT photometry (see their parameter estimates in Table 3.2). The basic idea of those authors is that the photometric observations provide clues on the

activity of CoRoT-7 , and because they are simultaneous to the HARPS observations, this information can be used to correct the activity-induced RV variations.

Basically, H14 used a GP to interpolate the photometric observations and then calculated the corresponding RV variations using the FF′ method of Aigrain et al. (2012). They found that this signal alone could not reproduce the full RV activity signal, and included a second GP, with a fixed covariance matrix, in the analysis of the RVs.

We took this approach one step further by trying to determine if the RV observations *alone* contained enough information to reproduce the signatures of both the planets and activity in a self-consistent manner. We analysed the full set of observations, with the RVs from 2012 and from the earlier campaign, in 2008-2009. In total, the star was observed with HARPS 177 times. The average error bar on these measurements is $2\,\mathrm{m\cdot s^{-1}}$ and the RV dispersion is $10\,\mathrm{m\cdot s^{-1}}$, over the complete 3 year timespan. Note that the RV error bars include photon and instrumental noise, being largely dominated by the former. More importantly, they do not contain a contribution from stellar activity. A plot of these observations is in Fig. 3.12, together with their generalised Lomb-Scargle periodogram.

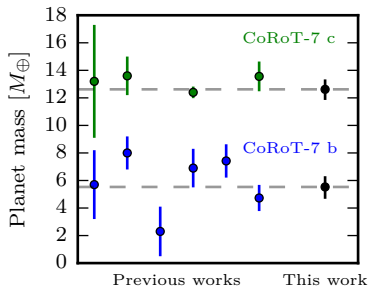


Figure 3.11: Planet mass estimates for CoRoT-7 b and CoRoT-7 c, from different analyses. Our estimates are presented in Table 3.3.

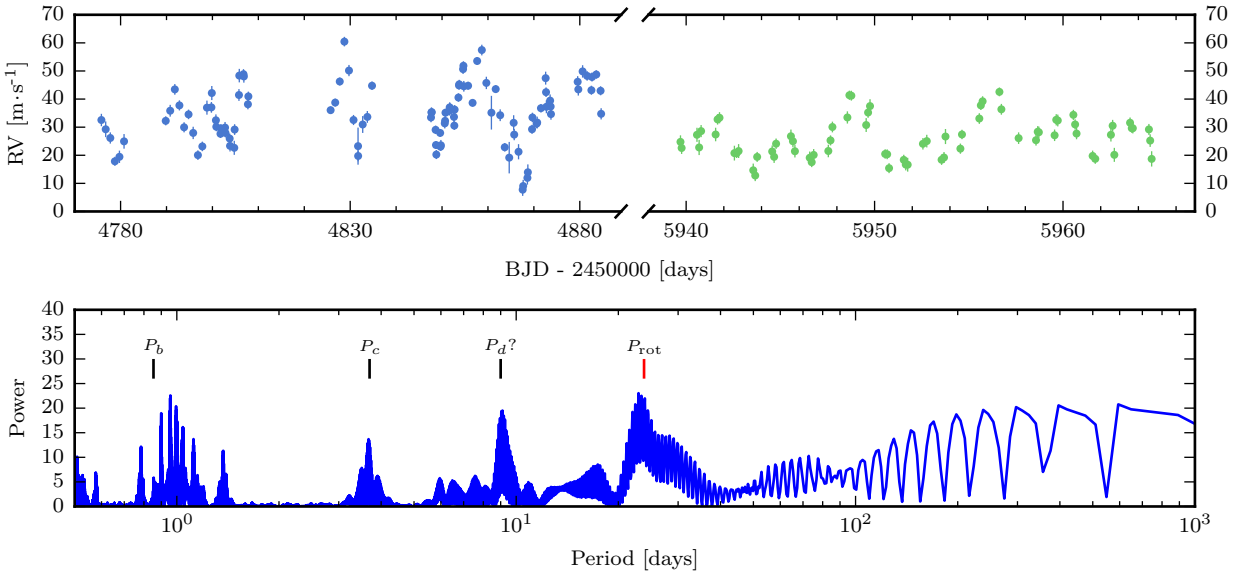


Figure 3.12: HARPS RV measurements of CoRoT-7 , for the two runs of about 4 months in 2008 − 2009 and ∼1 month in 2012. The bottom panel shows the generalised Lomb-Scargle periodogram of these observations, with an indication of the orbital periods of the planets and the rotation period.

The stellar rotation period is seen clearly in the periodogram, demonstrating that stellar activity is playing a very important role in the RV variations. Peaks at the periods of CoRoT-7 c and the proposed CoRoT-7 d are also clear, but the 1-day aliases mask the peak at the orbital period of CoRoT-7 b.

We ran our algorithm on this set of RVs, using the priors from Tables 2.1 and 3.1, and setting $N_{p,\max} = 10$. We obtained a total of

16 248 samples from the joint posterior distribution of all the parameters. With these samples, we calculated a value for the evidence of $\log(p(D \mid \mathcal{I})) = -530.9$. The resulting marginal posterior distribution for N_p is shown in Fig. 3.13. By applying the posterior ratio criterion, we choose $N_p = 2$ as the number of planets that are confidently detected from this dataset.

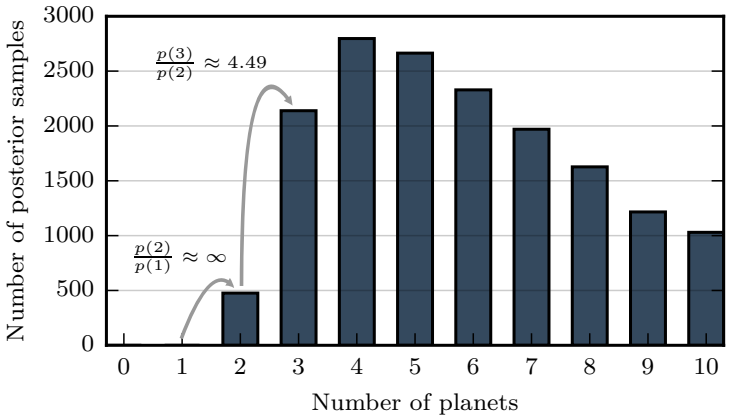


Figure 3.13: Posterior distribution for the number of planets N_p. The counts are number of posterior samples in models with a given number of planets. The two ratios of probabilities between models with 1, 2 and 3 planets are highlighted; note that no posterior samples were obtained with $N_p = 0$ and $N_p = 1$.

The joint posterior distributions for the orbital periods, semi-amplitudes and eccentricities of the signals are shown in Fig. 3.14, in which all the samples were combined. The figure shows an histogram of the posterior samples, where the colormap represents bin counts and is set in a logarithmic scale.

The two detected planets are seen as overdensity regions at $P_b = 0.85$ days and $P_c = 3.69$ days. Their amplitudes and eccentricities are well constrained. There is a clear posterior peak around 2 days with amplitude and eccentricity mostly unconstrained. The posterior also shows a smaller peak at 9 days, the period reported by Hatzes et al. (2010) as a possible third planet (see also Tuomi et al., 2014a).

The 2-day peak in the posterior for the orbital periods is currently lacking an explanation. Previous analysis of the RVs of CoRoT-7 did not mention any signal at this period. Whether this is an artifact of the GP activity correction or the result of the time sampling of the observations, is still unclear. In any case, do note that the colours of Fig. 3.14 are in a logarithmic scale and therefore the importance of the 2-day peak seems visually exaggerated.

Marginal and joint posterior distributions for the parameters of the GP and the extra white noise parameter, s, are shown in Fig. 3.15. The posterior for η_3 provides a constraint on CoRoT-7's rotation period, obtained exclusively from the RVs. Our inferred value for the stellar rotation period, of $22.30^{+10.08}_{-6.11}$ days has a large uncertainty but is in agreement with earlier estimates which used the CoRoT lightcurve

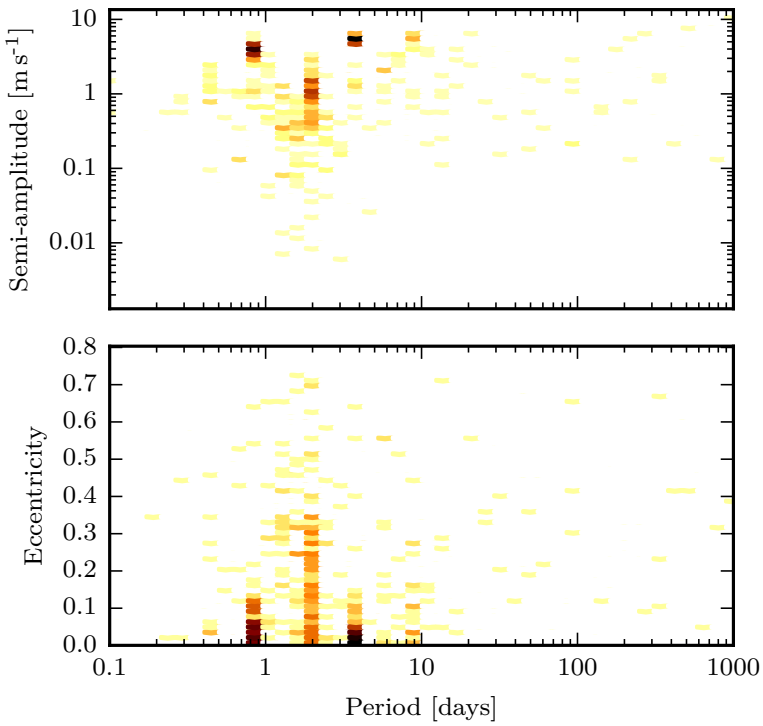


Figure 3.14: Joint posterior distribution for the semi-amplitudes (top panel) and the eccentricities (bottom panel) together with the orbital periods of the Keplerian signals.

(Léger et al. 2009; Lanza et al. 2010; H14).

Also interesting is the joint behaviour of η_2, η_3 and η_4. For higher values of η_4, the periodic component of the covariance function loses importance relative to the decaying component, η_2 gets smaller and η_3 becomes unconstrained. In this case the GP acts as an interpolator and smooths the RVs on a time scale of $\eta_2 \approx 3$ days.

But when η_4 is of order unity (meaning the periodic component is present), the decaying time scale is higher and η_3 is constrained around 22 days. The values of η_2 in this situation (around 20-30 days) are closer to the stellar rotation period and are also consistent with the average lifetime of active regions measured in the CoRoT 2012 photometry (20.6 $\pm$ 2.5 days, H14). A physical interpretation for this parameter, as the time scale for growth and decay of active regions, is then plausible.

Considering only the posterior samples with $N_p = 2$, Table 3.3 lists the median values of some orbital parameters for the two planets, and the maximum likelihood RV curves are shown in Fig. 3.16. Our estimates for the orbital parameters are in agreement, within the uncertainties, with the ones obtained by H14.

One of the most important points in our analysis is the use of RV data only. We found that we could recover the planetary and activity signals without considering photometric observations, any information from transit detections, or auxiliary activity indicators.

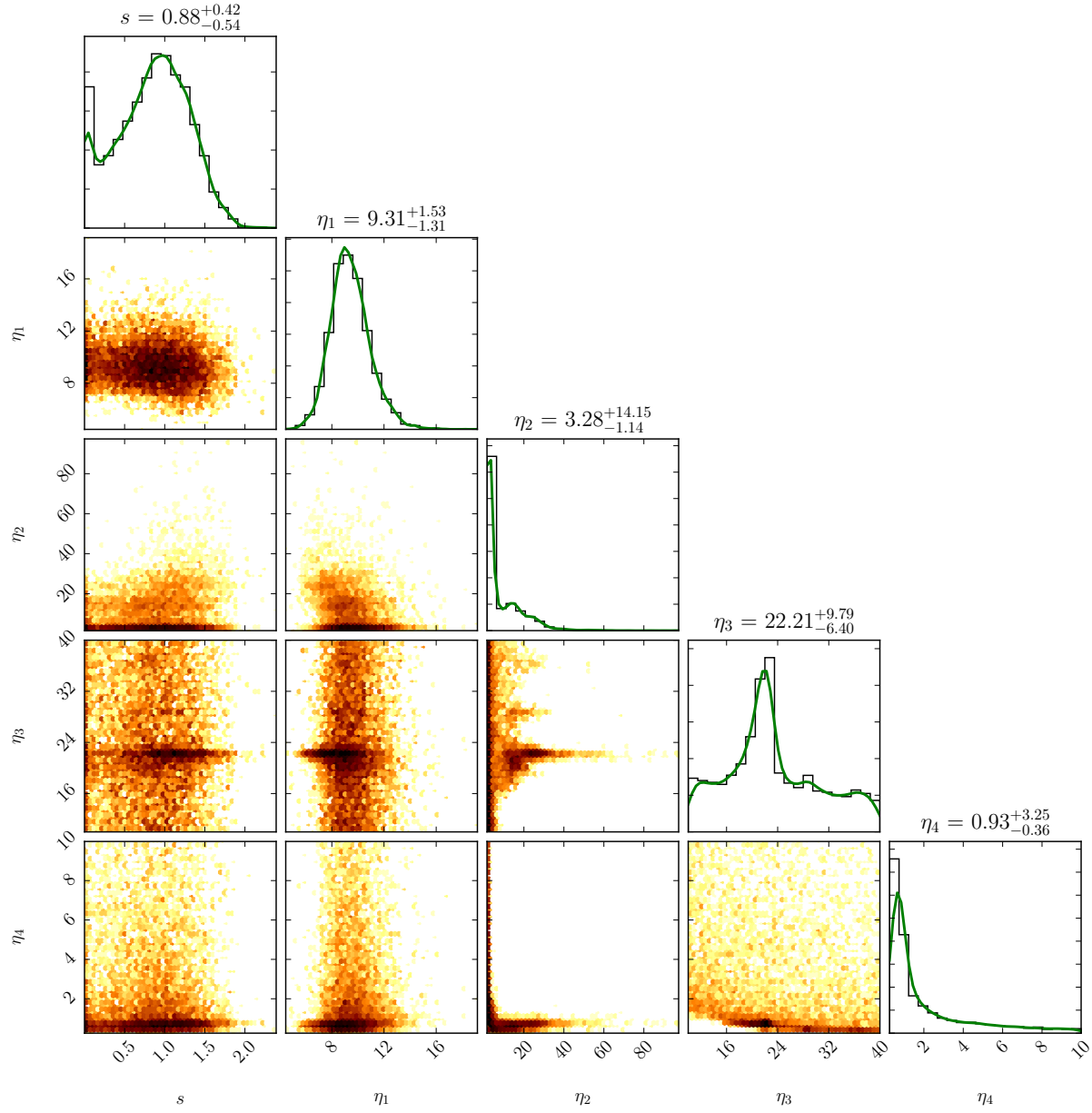


Figure 3.15: Marginalised 1- and 2-D posterior distributions for the parameters of the GP and the extra white noise. Samples for all values of N_p were combined. The titles above each column show the median of the posterior and the uncertainties calculated from the 16% and 84% quantiles. The solid lines are kernel density estimations of the marginal distributions. Adapted from Faria et al. (2016a).

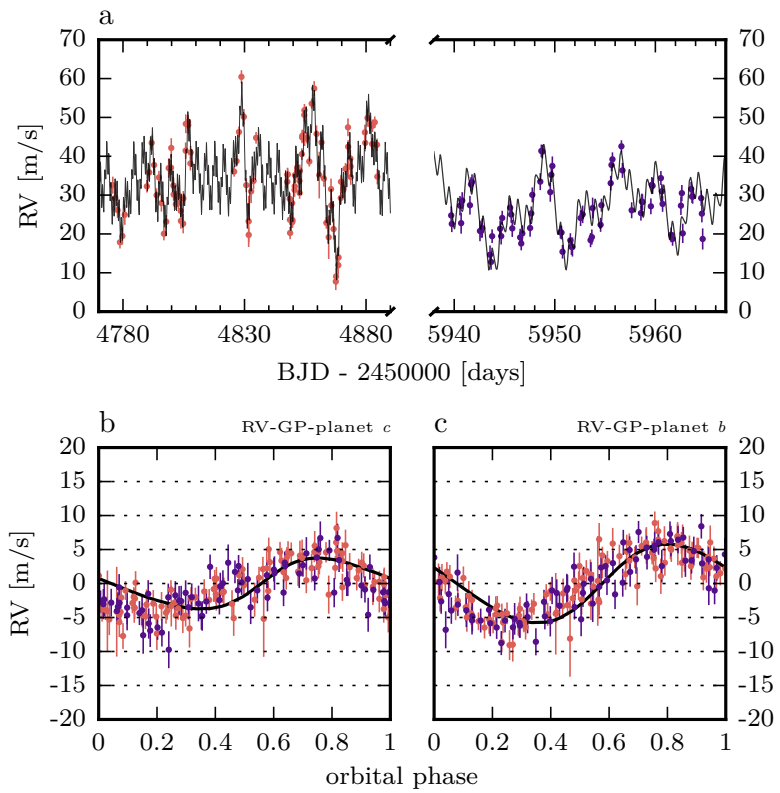


Figure 3.16: Panel a: RV measurements of CoRoT-7 from 2009 and 2012 and the two-planet best fit model (black curve). Note the different scales in the x axis of the left and right parts of the plot. At the bottom, panels b and c show the phased RV curves after subtracting each planet signal and the GP.

The GP provides a flexible and accommodating model for the activity-induced signals, allowing us to infer the planetary masses and orbital parameters with more realistic uncertainties.

This obviously does not mean that simultaneous photometry, the CCF activity indicators or any other information, are useless. If some of this information is available it can, and should, be used. For example, we could easily justify a more restrictive prior for η_3 based on photometric observations. Or we could have set a prior for η_1 based on the expected RV variations calculated with the FF′ method. In addition, any knowledge of the star that would help to constrain the GP kernel would be incredibly valuable.

I should also note two important properties of the CoRoT-7 system, which made it ideal for this analysis. First, the amplitudes of the planet signals are much higher than the mean error bar of the HARPS observations, regardless of the stellar activity contamination. Second, the time sampling of the observations is almost ideal for the detection of short-period planets, and is very difficult to obtain as part of a typical RV survey.

Of course, the fact that this model works well for CoRoT-7 does not necessarily mean it is optimal in modelling activity-induced RV variations. Further tests should be (and have been – see Chapter 4) undertaken to assess the limits of applicability of this method, either using other well-studied datasets of active hosts or simulated datasets.

Nevertheless, we were able to separate planetary and activity signals in the very difficult case of CoRoT-7, while making many fewer assumptions than previous studies. It is then reasonable to assume that this method can deliver good results on other cases, and that it should be used in other planetary candidates.

	units	this work	H14
P_b†	days	$0.85424^{+0.00071}_{-0.00126}$	$0.85359165 \pm 5\times10^{-8}$
K_b	$m{\cdot}s^{-1}$	$3.97^{+0.62}_{-0.55}$	3.42 ± 0.66
e_b		$0.045^{+0.053}_{-0.027}$	0.12 ± 0.07
m_b	$M_\oplus$	$5.53^{+0.86}_{-0.78}$	4.73 ± 0.95
P_c	days	$3.69686^{+0.00036}_{-0.00026}$	3.70 ± 0.02
K_c	$m{\cdot}s^{-1}$	$5.55^{+0.34}_{-0.31}$	6.01 ± 0.47
e_c		$0.026^{+0.033}_{-0.017}$	0.12 ± 0.06
m_c	$M_\oplus$	$12.62^{+0.77}_{-0.72}$	13.56 ± 1.08
η_3 ‡	days	$22.50^{+10.56}_{-6.19}$	23.81 ± 0.03

Table 3.3: Estimates for the parameters of the CoRoT-7 system, from our work and from H14. We consider all models that have $N_p = 2$ and show the marginal posterior medians together with the 16% and 84% quantiles.
The following notes apply to some of the estimates from H14:
† A Gaussian prior was used, centered at 0.85359165 days and with a standard deviation of 5.6×10^{-7} days.
‡ The estimate for the rotation period was derived from the CoRoT lightcurve.

4

The metal-poor sample

Being a probe of the metal content of proto-planetary disks, stellar metallicity has long been recognized as an important ingredient for planet formation. As was discussed in Chapter 1, giant planets have been found to be more frequent around metal-rich stars. But in the metal-poor regime, the frequency of planets – especially low-mass planets – and the way in which it depends (if at all) on metallicity are still subject of debate.

In this Chapter, I present and analyse the data from an HARPS RV survey focused on moderately metal-poor stars. I start with a brief historical account of some RV surveys that motivated our metal-poor program. Afterwards, a description of our sample of metal-poor solar-like stars and the derivation of their fundamental stellar parameters is shown. Then I discuss the analysis of interesting individual stars, the ensemble analysis of all stars, and finally the determination of the occurrence rate of small planets in the complete sample.

4.1 METALLICITY IN RADIAL-VELOCITY SURVEYS

Stellar metallicity has had an impact in RV surveys. Because of the giant-planet – metallicity correlation (see section 1.2) and in order to increase the likelihood of finding planets,[1] a number of surveys were specifically dedicated to metal-rich samples, such as those described in Tinney et al. (2002), Fischer et al. (2005), Da Silva et al. (2006), and Melo et al. (2007).

But a few giant planets were also found around metal-poor stars, as reported by Setiawan et al. (2003), Mayor et al. (2004), and Cochran et al. (2007), for example.[2] A few programs then concentrated in searching for planets around low-metallicity stars. Two of these used the Keck and HET telescopes (Sozzetti et al., 2006; Cochran and Endl, 2008) and a third one was part of the HARPS GTO program (Mayor et al., 2003), as described in Santos et al. (2007). The first detections from this HARPS metal-poor survey were giant planets (e.g. Santos et al., 2010b) orbiting some of the stars which would eventually be included in our sample (see section 4.3.2).

The planets discovered in these surveys show that giant planet formation is not completely inhibited around stars in the metal-poor

[1] In particular, by surveying more metal-rich stars one is more likely to detect short-period giant planet companions, also increasing the chance of detecting a transit.

[2] This last discovery is actually quite remarkable in that two Jupiter-mass planets were found orbiting HD 155358, with a metallicity [Fe/H] $= -0.68$, and thus, only 21% of the heavy-element content of the Sun. The reported two-planet system of HD 47536, from Setiawan et al. (2003, 2008), would also be remarkable, but new data has disproved the second companion (Soto et al., 2015).

regime. They can represent either the metal-poor tail of the giant planets formed by core-accretion, or a different population of planets formed as a result of disk-instability processes (see Maldonado and Villaver, 2017; Santos et al., 2017, and references therein).

The observing strategy and the relatively low RV precision of the low-metallicity surveys mentioned above, meant that these were not adequate to detect lower mass planets, such as Neptunes and super-Earths. In order to probe this new region of parameter space – low-mass planets around low-metallicity stars – a dedicated HARPS program started in October 2008, focused on a subsample of the HARPS GTO targets.

The objective of this program was to constrain the presence of low-mass planets around metal-poor stars. But with the first observations, there were no clear signals caused by low-mass planets.[3] The initial program was then continued for an extra 80 nights over the three upcoming years (starting in October 2012), within the ESO Large Program with reference 190.*C*-0027. The 109 stars from this Large Program (LP) constitute the complete sample I analyse here.

[3] The detection of two planets around HD 41248 (Jenkins et al., 2013) was published around this period, but the history of this system did not stop there, as will be discussed in Section 4.3.4.

Santos et al. (2014) described the complete sample and the selection criteria: from the GTO sample we selected bright ($V < 9.5$), chromospherically quiet ($\log R'_{\rm HK} < -4.8$) and low-metallicity stars (Fe/H< -0.4; as derived from the HARPS CCF with the method of Santos et al. 2002) that were not known to host any low-mass planets. This defined a sample of 105 metal-deficient FGK stars (including also one star classified as an M dwarf) which are suitable targets for detecting very low mass planets. To these we added the three planet hosts presented in Santos et al. (2011) as well as a fourth long-period planet-hosting candidate (see section 4.3.2).

Following this LP, another 51 nights were granted to a new program (196.*C*-0042) which focused on a smaller sample of only 22 stars, out of the original 109. For this program we selected stars which were particularly inactive, with $\log R'_{\rm HK} < -4.9$, rotate slowly ($v \sin i < 3\,{\rm km\cdot s^{-1}}$), had already at least 10 RV measurements, and are brighter than $V = 9.0$. With HARPS, this should allow for a photon noise precision better than $1\,{\rm m\cdot s^{-1}}$ after a 20 minute exposure. We have also chosen only G and K stars (excluding late F stars that were in the initial sample) for which the oscillation and granulation noise is expected to be lower (Bouchy et al., 2005; Dumusque et al., 2011c). To alleviate complicated selection effects, no star was added because it had a candidate planetary signal. This means that we can include these new RV observations in the statistical analysis of the full sample.

This program is still ongoing (at September 2017) with the goal of improving the detection limits of the target stars (see section 4.4.1). Some of the stars now have precise HARPS RV observations dating

back 11 years, although early observations usually show higher scatter and are very dispersed in time.

The distribution of the number of RV measurements is shown in Fig. 4.1. Many stars have a few tens of measurements and a few were followed quite intensively with a few hundred points. The scale in the figure shows the minimum, mean and maximum number of observations, and Table C.2 in Appendix C, shows more detailed information for the sample stars: the number of nights each star was observed, the total number (not nightly binned) of observations, the mean error bar $\bar{\sigma}_k$, the weighted standard deviation of the radial velocities $s_{\rm RV}$ and the time baseline of the observations.

The ratio $s_{\rm RV}/\bar{\sigma}_i$ gives an indication of whether the radial velocities vary in excess of the internal errors, and by how much. This ratio is sometimes considered in an F-test for RV variability (e.g. Zechmeister et al., 2009). In Fig. 4.2 we plot these two quantities as a function of the stellar metallicity (see the next section for a discussion on how we derive the stellar metallicities). A few targets stand out, showing large RV variations due to the presence of giant planets, long-term trends or activity signals. Otherwise we note a small decrease in the RV precision for lower metallicities. As discussed in Section 2.2.4, this is associated with the lower number of lines and their smaller depths.

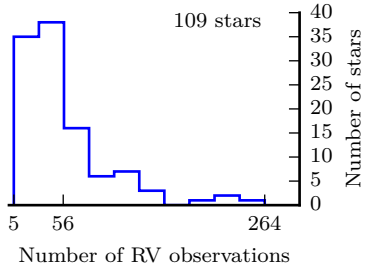


Figure 4.1: Distribution of the number of RV observations for the stars in the sample, highlighting the minimum, mean, and maximum values.

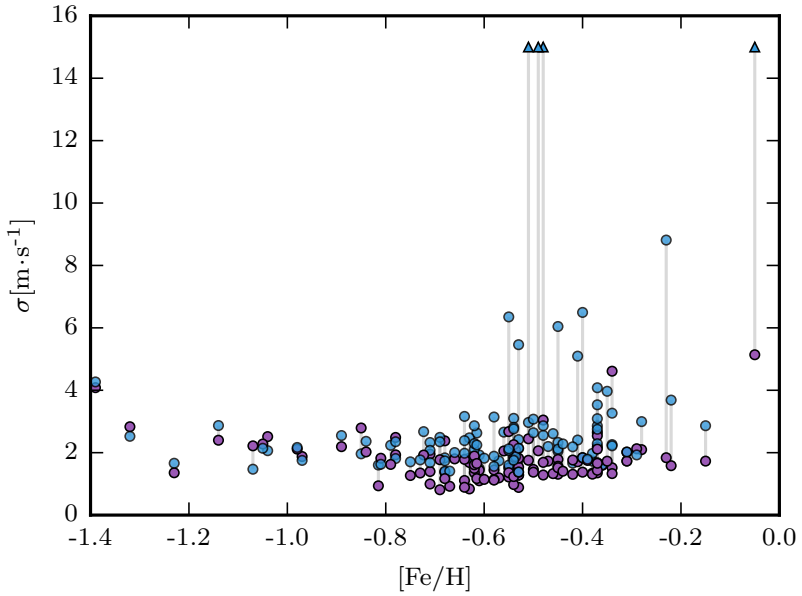


Figure 4.2: The weighted standard deviation of the radial velocities $s_{\rm RV}$ (blue circles, changed to triangles when clipped to $15\,{\rm m\cdot s^{-1}}$) and the mean error bar $\bar{\sigma}_i$ (purple circles) for each star as a function of its metallicity. For stars with observations after the HARPS fibers changed, we estimated and corrected for the RV offset using a weighted average (see section 4.3.1).

4.2 STELLAR PARAMETERS

The sample of stars in the LP includes 109 targets with spectral types F, G, K and one classified as an M dwarf. As mentioned before, these targets were observed with HARPS during several runs. The

individual spectra have a resolution $R \sim 110000$ and signal-to-noise ratios (SNR) that vary from 6.8 to 459. To achieve a higher SNR for the derivation of stellar parameters, the spectra for each star (those with SNR > 20) were co-added, after correcting for their radial velocity.

To derive spectroscopic parameters – the effective temperature T_{eff}, surface gravity $\log g$, metallicity [Fe/H] and microturbulence ξ – for the sample, we use the method introduced in Santos et al. (2004b), which relies on a differential analysis with respect to the Sun. This is only appropriate for F, G, and K stars with temperatures ranging from $\sim$4500 K to $\sim$6400 K. For HD 304636, the only star in our sample classified with spectral type M0, we used the method presented in Neves et al. (2014) to derive its effective temperature and metallicity.

The procedure for stellar parameter determination has been described and used in a series of publications: Sousa et al. (2006, 2008, 2011a) and more recently in Sousa (2014). It is based on measuring the equivalent widths (EW) of Fe I and Fe II weak lines, and on computing the abundance of each line, assuming local thermodynamic equilibrium. For the automatic measurement of the EWs we use the ARES code (Sousa et al., 2007), and for abundance determination, the MOOG code (Sneden et al., 2012). The procedure then follows these steps (see Santos, 2002):

- From an initial guess of the stellar atmospheric parameters T_{eff}, $\log g$, [Fe/H], ξ, an atmosphere model is generated and introduced in MOOG, together with the measured EWs for the iron lines. MOOG then computes the iron abundance that fits each spectral line under the considered atmosphere model.
- Microturbulence: if the microturbulence value, ξ, is the correct one, all lines should yield the same abundance $\epsilon(\mathrm{FeI})$, regardless of their EW. Therefore, we test if the slope of the relation $\log \epsilon(\mathrm{FeI})$ vs. EW is close to zero and otherwise restart the procedure with a different value of microturbulence (the other parameters are kept fixed).
- Temperature: if the effective temperature value is the correct one, all lines should yield the same abundance, independent of their excitation potential, χ (see for example Gray, 2005). We test if the slope of the relation $\log \epsilon(\mathrm{FeI})$ vs. χ is close to zero and otherwise restart the procedure with a different value of effective temperature (the other parameters are kept fixed).
- Surface gravity: the average abundances obtained from the FeI and FeII lines are compared and the $\log g$ value is changed until they are similar (the other parameters are kept fixed).
- When the three parameters T_{eff}, $\log g$, and ξ are determined, the resulting Iron abundances are used to restart the procedure. This is done until all 4 parameters (including [Fe/H]) converge.

The internal uncertainties in the atmospheric parameters are of the order of 50 K, 0.15 dex, 0.1 dex and 0.06 dex for T_{eff}, $\log g$, ξ, and [Fe/H], respectively. Note that the Iron lines used to derive atmospheric parameters should be chosen so that the EWs can be measured accurately. Ideally, they should not be blended or saturated. Different line lists exist in the literature (e.g. Tsantaki et al., 2013), for either general applications or for specific types of stars. For the stars in the LP we used the line list of Sousa et al. (2011b).

The stellar masses were estimated following Santos et al. (2013), in order to keep uniformity with the SWEET-Cat catalogue. The values of temperature, surface gravity and metallicity derived before were used as input to the calibration presented in Torres et al. (2010). We also consider the correction mentioned by Santos et al. (2013) to obtain comparable masses to those derived using Padova isochrones (da Silva et al., 2006). To derive uncertainties for a stellar mass M_*, we use a simple Monte Carlo analysis: 10 000 random values of effective temperature, surface gravity, and metallicity are drawn assuming a Gaussian distribution with their uncertainties, and then used as input to the Torres et al. (2010) calibration. From the resulting distribution of masses, we quote the mean and the standard deviation. The intrinsic error from the Torres et al. (2010) calibration ($\sigma_{\log M_*} = 0.027$) is then quadratically added to get the final uncertainty.

Fig. 4.3 shows the distributions of metallicity, T_{eff}, $\log g$ and stellar mass, with the minimum, mean and maximum values for each quantity. Table C.1, in Appendix C, lists all the derived parameters for the 109 stars. The resulting parameters show that our sample is made of solar-type, moderately metal-poor dwarfs. Note that some targets have metallicities higher than the -0.4 threshold used to construct the sample. This is because that threshold was evaluated against the value of [Fe/H] derived from the CCF (Santos et al., 2002), and not from a detailed spectroscopic analysis.

In comparison with the HARPS GTO sample presented in Mayor et al. (2011), which has a mean [Fe/H] of -0.10, it is clear that our sample is complementary to that one but focuses on substantially more metal-poor stars. This metallicity distribution is very important, as our sample has been built to answer the specific question of the frequency of planets in the metal-poor regime.

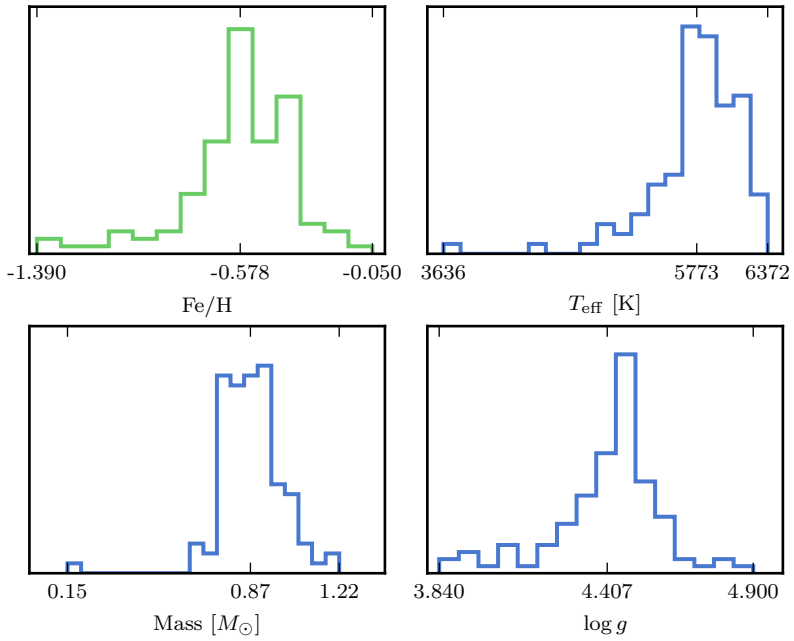


Figure 4.3: Distributions of metallicity, effective temperature, stellar mass and surface gravity for the 109 stars in our sample. The axes show the minimum, mean and maximum value of each distribution.

At this point I also introduce the *starfish diagram* (Konstantopoulos, 2015) in Fig. 4.4, which allows for a quick visual placement of an individual star within the sample distributions, so that we can more easily identify common trends for groups of stars. The diagram shows the values of effective temperature, magnitude in the V band, metallicity, surface gravity and the number of RV observations available for each star. These values increase from the center of the diagram outwards. The complete sample is shown as a black distribution and any individual star in orange.

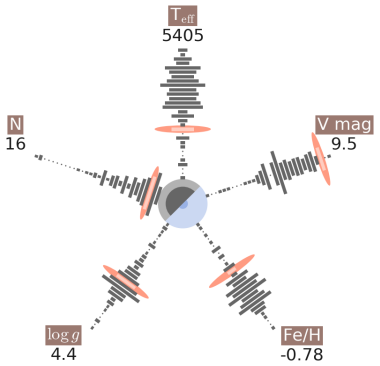


Figure 4.4: Starfish diagram for an example star in our sample. The diagram shows the effective temperature, V magnitude, metallicity, surface gravity and number of RV observations for a given star, placed within the sample distributions of each parameter.

4.3 FINDING PLANETS IN THE METAL-POOR SAMPLE

In this Section, I will analyse the complete set of RV timeseries for all the 109 stars in our sample. First, I start by discussing the inclusion in our model of two extra parameters related to a linear trend and a RV offset after the HARPS fibers change. A few interesting individual stars are then the subjects of detailed analyses in order to uncover or rule out the presence of planets. Afterwards, results from the ensemble analysis of the full sample of stars in the LP are discussed.

4.3.1 *Long-term trends and fiber offset*

The model for RV variations presented in Chapters 2 and 3 is almost readily applicable to the analysis of all the LP targets. However, we must first consider the inclusion of two extra parameters to take into account the presence of long-term linear trends and the RV offset

introduced by the change in the HARPS fibers.

A visual inspection of the RV timeseries in our sample, shows the presence of clear long-term trends for a few stars, namely HD 11397, HD 144589, HD 197083 and HD 93351 (see the RV timeseries in pages 169, 172, 175, and 184). When this is the case, if we use the RV model developed earlier, we will see a typical posterior distribution for N_p, indicating the presence of one planet with a very long orbital period. Because of the hyper-prior for P (see Table 2.1), which states that orbital periods of two or more planets should be within a few orders of magnitude, we can be biased if we try to estimate the parameters of any additional companions.

To deal with this, it is common to include a linear drift parameter in the analysis of RV timeseries (e.g. Feroz and Hobson, 2014). This drift can account for RV variations caused by the presence of a distant companion, which cannot be accurately characterized within the time span of the observations. Of course, including this drift raises some issues: Should a quadratic term also be considered? Should this parameter be added in the analysis of every star, or only when the trend is clearly identified visually?

It is important to notice that only a large enough trend would bias substantially the estimation of the other parameters, and such a trend will be almost certainly clearly visible in the RV timeseries. Therefore, I have decided to include a linear drift in the model only for those stars in which the trend can be identified visually or if the posterior samples (from a model without the drift) show very long period signals, larger that the timespan.

In practice, we only need to consider an extra parameter, tr, and change Eq. (2.64) into

$$v_r(t) = v_{\rm sys} + tr\,(t - t_1) + \sum_i K_i\{\cos\left[\omega_i + f_i(t)\right] + e_i\cos\omega_i\}\,, \quad (4.1)$$

where i indexes the planets and t_1 is the time of the first observation. The prior for tr could be determined in a number of ways. As the simplest option, we consider the following uniform distribution

$$tr \sim \mathcal{U}\left(-\frac{\max v_{\rm obs}(t_k) - \min v_{\rm obs}(t_k)}{\max t_k - \min t_k}, +\frac{\max v_{\rm obs}(t_k) - \min v_{\rm obs}(t_k)}{\max t_k - \min t_k}\right) \quad (4.2)$$

with the limits efficiently bracketing the possible linear trends supported by the data (see Fig. 4.5). Note that this prior depends explicitly on the available data, but it can still be regarded as uninformative.

In practical terms, our choice of including only a linear trend instead of, e.g., a quadratic trend, does not have a serious impact on any of the results. This is because the prior for the orbital periods (see Table 2.1) still allows for very large values of P (much larger

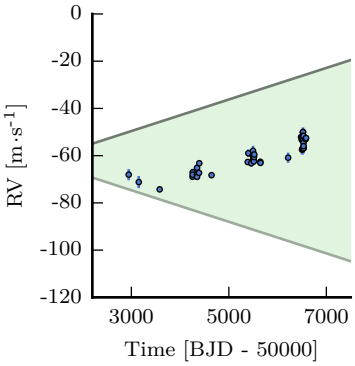


Figure 4.5: Prior region allowed for a linear trend, using the data of HD 197083 as an example. The two grey lines are defined by the priors for $v_{\rm sys}$ and tr.

than the timespan of the data), if indeed these values are supported by the data. Therefore, our choice is merely a compromise between simplicity and interpretability.

* ◯ *

As was mentioned in section 2.2.5, new octagonal optical fibers were installed in HARPS in June 2015. Changing the fibers has an impact on how the image of the stellar lines is formed on the detector. Therefore, a different fiber set corresponds in practice to a different instrument. The change in the fibers introduced an RV offset of several $\mathrm{m{\cdot}s^{-1}}$, correlated with the width of the stellar lines (the FWHM), as described by Lo Curto et al. (2015). Many stars in the LP have observations after this date, which corresponds to BJD 57170 (see Fig. 4.6). In fact, a number of stars from our sample were actually included in Lo Curto et al. (2015) as standard stars[4] and used for the estimation of the RV offset. Therefore, in order to take into account this RV offset, a new parameter δ was included in the model such that Eq. (4.1) becomes

[4] Whether they can be considered as "standard" or not is subject to debate.

$$v_r(t) = \begin{cases} v_{\mathrm{sys}} + tr\,(t - t_1) + \sum_i K_i\{\cos\left[\omega_i + f_i(t)\right] + e_i \cos\omega_i\} & \text{for } t < 57170 \text{ MJD} \\ v_{\mathrm{sys}} + tr\,(t - t_1) + \sum_i K_i\{\cos\left[\omega_i + f_i(t)\right] + e_i \cos\omega_i\} + \delta & \text{for } t > 57170 \text{ MJD} \end{cases} \tag{4.3}$$

and we used as prior for δ a uniform distribution (in $\mathrm{m{\cdot}s^{-1}}$)

$$\delta \sim \mathcal{U}\ (0, 30)\ , \tag{4.4}$$

where the upper limit of $30\,\mathrm{m{\cdot}s^{-1}}$ was set by visual inspection of the time series. The offset is expected to be always positive for the range of FWHM of the stars in our sample. Note also that even though we are only extending our model for RV variations, other quantities show offsets as well, namely the FWHM itself, the CCF contrast and even the $\log R'_{\mathrm{HK}}$ index. A complete study of these offsets and the full effect of the fiber change is out of the scope of this Thesis.

These two new parameters, tr and δ, should be added to the probabilistic graphical model of Fig. 3.10, and to the vector $\boldsymbol{\theta}$. Like v_{sys}, they are connected only to $v_{\mathrm{model}}(\boldsymbol{\theta}, t)$.

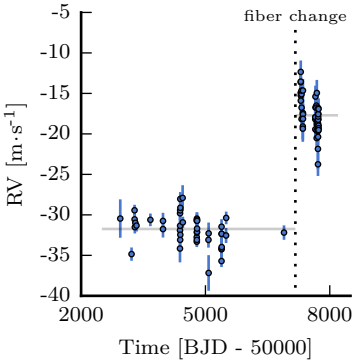


Figure 4.6: Example of a RV time-series (of HD 967), showing the offset introduced by changing the HARPS fibers.

* ◯ *

To show a few particular examples of the issues discussed above, I present here the analysis of some targets that show clear long-term trends and have observations made after the fiber change. Fig. 4.7 shows the RVs of HD 11397, HD 144589, HD 197083 and HD 93351, together with the best fit linear regression line. These results were obtained with our RV model by fixing N_p to 0. The middle panels show the residuals from these fits and in the bottom panels the GLS

periodograms of the residuals and the 1% FAP line.

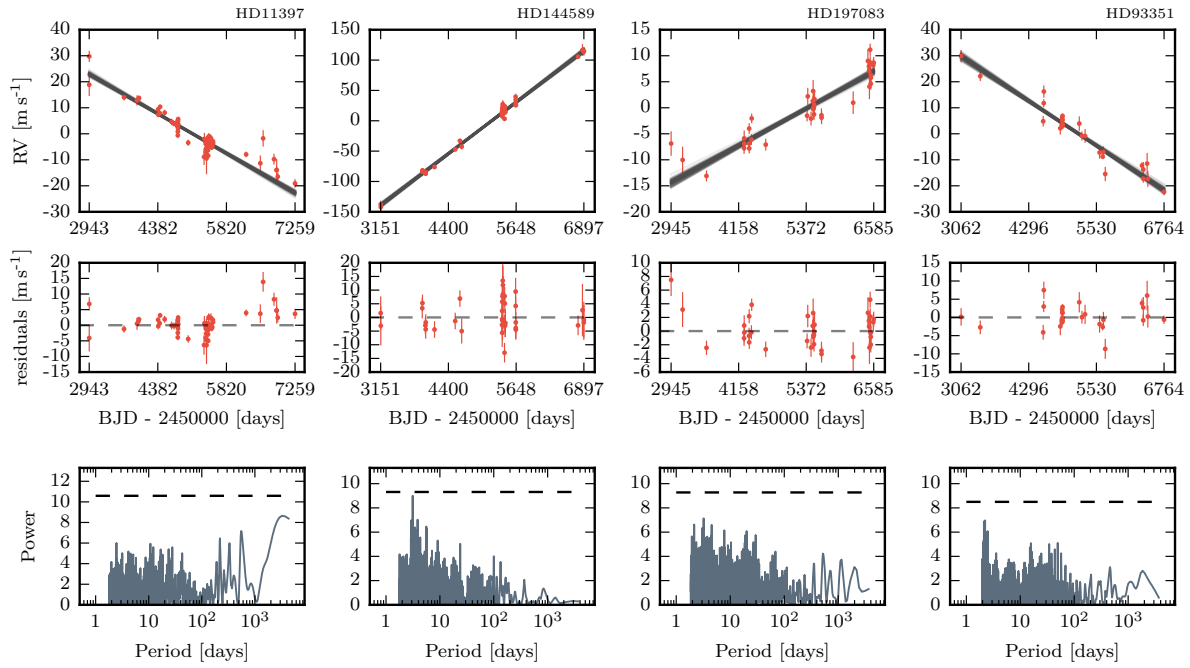


Figure 4.7: Initial analysis of stars with clear linear trends, with N_p fixed to 0. The top panels show the RVs and posterior predictive samples, the middle panels show the residuals of the fits and the bottom panels the GLS periodograms of the residuals.

After this initial analysis, one could in principle conclude from the periodograms that the RVs do not show any planetary signals. But now that we have extended our model to include the *tr* parameter, we can run the complete inference and obtain the posteriors for N_p. We set $N_{p,\max}$ to 3 and obtain 100 000 samples from the Diffusive Nested Sampling target distribution. The resulting posterior distributions for N_p are shown in Fig. 4.8 for the four stars. HD 11397 stands out, with the only posterior that shows evidence for the presence of one Keplerian signal, which we discuss below.

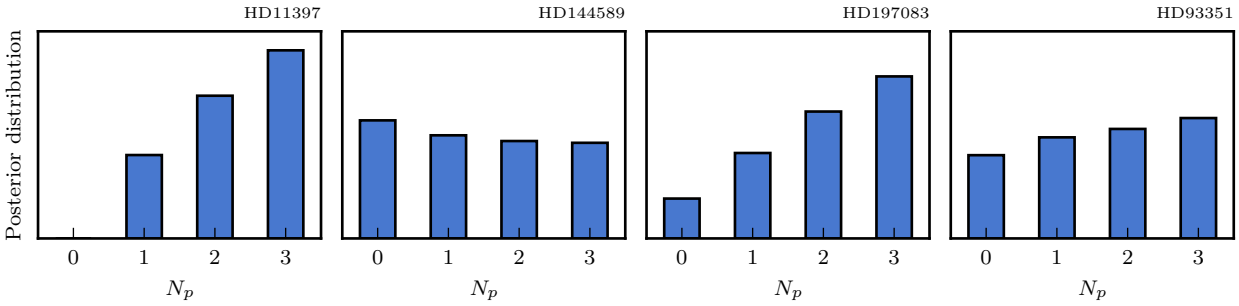


Figure 4.8: Posterior distributions for N_p for the four targets with clear long-term trends. The model used here includes the parameter *tr*.

Being a parameter in our model, we also infer the posterior distribution for *tr*, shown in Fig. 4.9 for the four stars. Assuming that the linear trends are caused by undetected, longer period companions, we can take into consideration the timespan of the RV observations to infer some properties of those companions.

In very basic terms, the presence of a linear trend over a timespan Δt can only be caused by a companion with orbital period $P > 2\,\Delta t$, if we assume a circular orbit. The estimated value of $tr \cdot \Delta t$ provides us with the minimum amplitude allowed for the companion's RV signal, so that the semi-amplitude is $K_{\rm min} = tr \cdot \Delta t/2$. Using the known stellar masses and assuming a circular orbit, we can then constrain the mass of the companion as a function of orbital period.

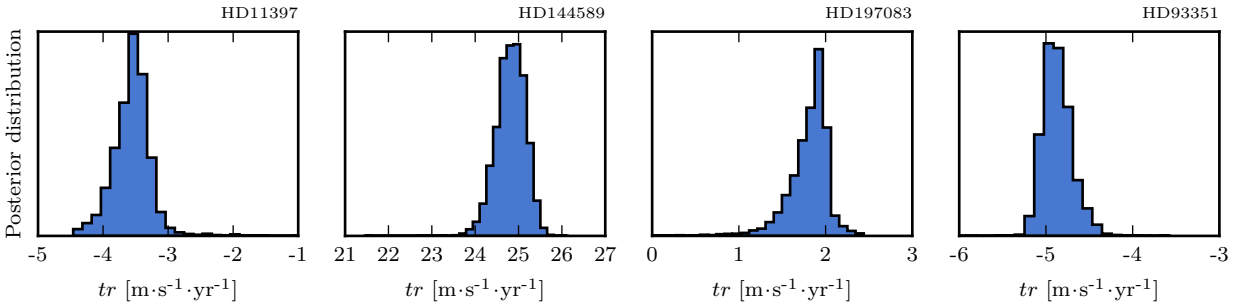


Figure 4.9: Posterior distributions for the slopes of the linear trends.

Another option, arguably simpler, it to actually run our RV model fixing N_p to 1, and not including the linear trend tr. This will force the linear trend to be modelled as a Keplerian function allowing us to estimate the posterior distribution for the orbital period of a companion, while removing the limitation of considering only circular orbits. Our tests show that this option gives compatible constraints to the simple calculations above. But it is clear that it provides the most information, at the expense of an increased computational cost.

Using the second approach, we obtained constraints for the orbital periods and masses of the long-period companions of HD 11397, HD 144589, HD 197083, and HD 93351. Listing the median value and the 68% quantiles of the posterior distributions for P and for the companion mass, the constraints are:

HD 11397: $P = 12460^{+6366}_{-3556}$ days and $m_p \sin i = 2.53^{+1.14}_{-0.69}\ M_{\rm Jup}$

HD 144589: $P = 29675^{+49041}_{-15669}$ days and $m_p \sin i = 86.18^{+233.47}_{-50.01}\ M_{\rm Jup}$

HD 197083: $P = 9614^{+9499}_{-3868}$ days and $m_p \sin i = 1.42^{+1.37}_{-0.61}\ M_{\rm Jup}$

HD 93351: $P = 19802^{+30378}_{-9058}$ days and $m_p \sin i = 7.83^{+19.45}_{-3.90}\ M_{\rm Jup}$

From these results we can see that the long-period companions to both HD 144589 and HD 93351 are most likely brown dwarfs, as their mass constraints places them above the deuterium burning limit (Spiegel et al., 2011). For HD 11397 and HD 197083 we obtain Jupiter-like planets on long periods. We stress that these constraints are based on an incomplete coverage of the companion orbits and therefore are subject to change with new data.

We can now go back and analyse the results of HD 11397, where the posterior for N_p, with the more general model including tr, suggested the presence of one Keplerian signal in addition to the long-term trend. Out of the 70 RV observations of this star, only 2 were made after the change in the HARPS fibers (see the timeseries in page 169). The detection of a significant signal is independent on the inclusion of these last observations, but they do change the posterior for the orbital period. One should note, however, that their inclusion means that the parameter δ is being constrained by only two RV points.

To improve the constraints on the orbital parameters, we ran our model with N_p fixed to 1, considering those two last RV points. The resulting posterior distribution for the orbital periods shows a clear peak close to 2000 days (Fig. 4.10), but solutions with longer periods, near the timespan of the data, are also probable. These two possible solutions are strongly correlated with the parameter δ, and therefore depend on the inclusion of the last two points. Fig. 4.11 shows joint posterior samples for δ and P. There are clearly two isolated modes in the posterior distribution, with the periods around 2000 days corresponding to generally larger values of δ.

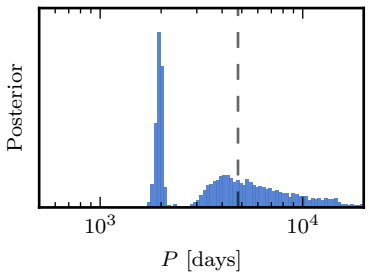


Figure 4.10: Posterior for the orbital period, from the analysis of HD 11397. The dashed line marks the timespan of the observations.

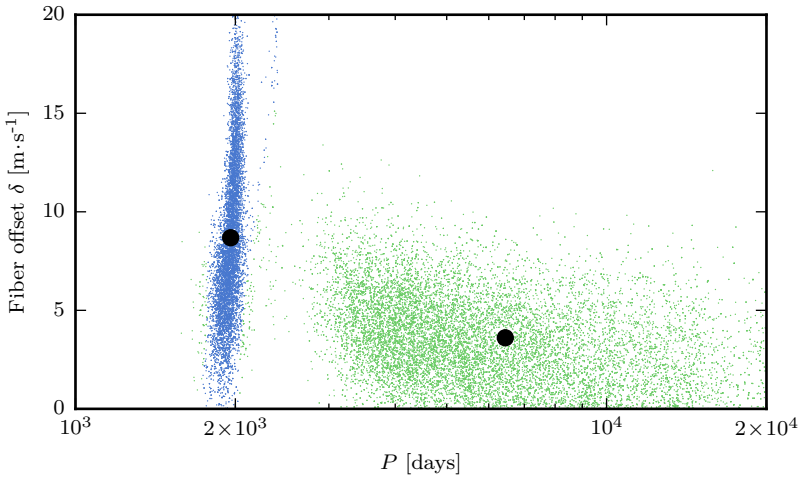


Figure 4.11: Samples from the joint posterior distribution for δ and P. Two separate clusters are identified and the black circles mark the means of the samples on those clusters.

According to the results of our analysis, the most probable value of the orbital period is around 2000 days. The values of δ that correspond to these solutions also agree better with the linear relation found by Lo Curto et al. (2015) between tr and the FWHM of the star. The mean value of the blue samples in Fig. 4.11 is $\delta = 8.68\,\mathrm{m\cdot s^{-1}}$, which is compatible with the average FWHM of $6.31\,\mathrm{km\cdot s^{-1}}$ for HD 11397.

Using only the posterior samples shown in blue in Fig. 4.11 (which were obtained after applying a cut in P), our estimates for the orbital parameters of HD 11397 b give a period $P = 1962.45^{+61.74}_{-75.41}$ days, a semi-amplitude $K = 3.87^{+3.29}_{-1.05}\,\mathrm{m\cdot s^{-1}}$, and a significant eccentricity $e = 0.46^{+0.21}_{-0.16}$. These parameters translate to a companion mass of

$m_p \sin i = 0.18^{+0.10}_{-0.05}\ M_{\rm Jup}$. In summary, the analysis of the RV observations of HD 11397 results in two long-period giant planets.

4.3.2 *Known giant planets*

A few of the stars included in the metal-poor sample are known hosts of giant planets, the rationale for their inclusion in this sample being the detection of possible lower mass companions. In this section, I analyse the RV observations available for HD 171028, HD 181720, HD 190984, and HD 107094. For reference, the orbital periods and masses of these planets, as announced by the discovery papers, are shown in Table 4.1. Starfish diagrams with some parameters for these stars are in Fig. 4.12 for a visual comparison. The derived metallicities for the four stars place them close to the average [Fe/H] of our sample.

	Period [days]	Mass [$M_{\rm Jup}$]
HD 171028 b	538 ± 2	1.83
HD 181720 b	956 ± 14	0.37
HD 190984 b	4885 ± 1600	3.1
HD 107094 b	1870 ± 34	4.5

Table 4.1: Orbital periods and minimum masses of the known giant planets, as derived by Santos et al. (2007, 2010b, 2011).

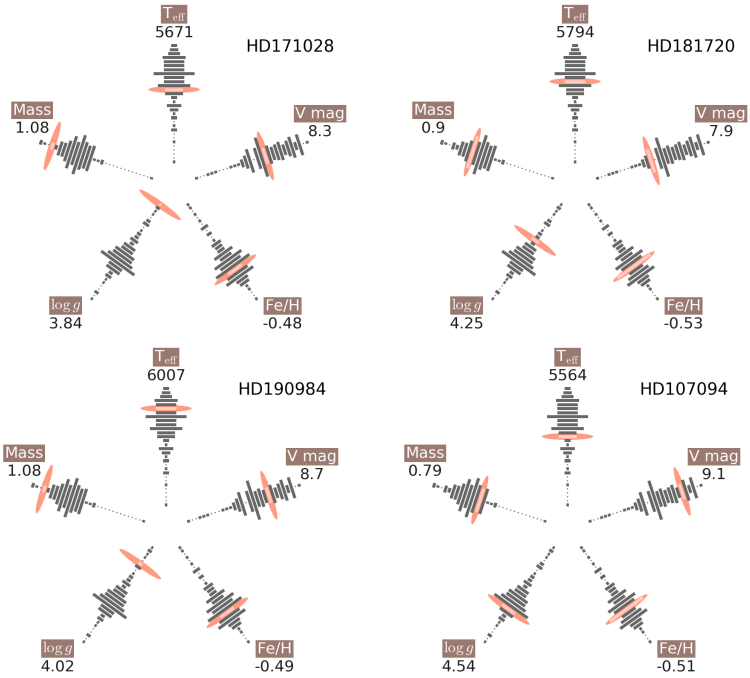


Figure 4.12: Stellar parameters for the stars in the LP sample known to host giant planets.

The derived parameters for HD 171028 make it a slightly evolved solar-type star, and are compatible with the G0 spectral type listed in the Simbad database. We obtain a slightly higher effective temperature $T_{\rm eff} = 5671 \pm 16$ K but compatible with the value reported by Santos et al. (2007). At the time of the discovery paper, HD 171028 was not listed in the Hipparcos catalogue, and only a rather inaccurate parallax of 9.1 ± 7.8 mas was available from the Tycho catalogue. Fortunately, the star was included in the first data release from GAIA (Lindegren et al., 2016) providing a new parallax of 8.86 ± 0.23 mas.

For HD 190984, Santos et al. (2010b) rederived the parallax because the measurement available from the Hipparcos catalogue ($\pi = 5.45 \pm 1.11$ mas; van Leeuwen, 2007) had a large uncertainty. They obtained

a value of 9.8 ± 2.6 mas, significantly different from that listed in the Hipparcos catalog. This star is also included in the GAIA DR1, now having a more precise value of $\pi = 6.84 \pm 0.25$ mas. HD 107094 is also included in GAIA's first data release, with a measurement of the parallax of 18.83 ± 0.28 mas. For HD 181720, we obtain compatible stellar parameters to those presented in Santos et al. (2010b).

From the HARPS spectra we derived average values for the $\log R'_{\rm HK}$ chromospheric activity index of -5.05, -4.99, -5.04, and -4.82 for HD 171028, HD 181720, HD 190984, and HD 107094, respectively. Using also the B-V colors, these correspond to rotation periods of 25.1, 16.7, 12.9, and 21.1 days, respectively, using the calibration of Noyes et al. (1984b). These values are shown in Table C.1.

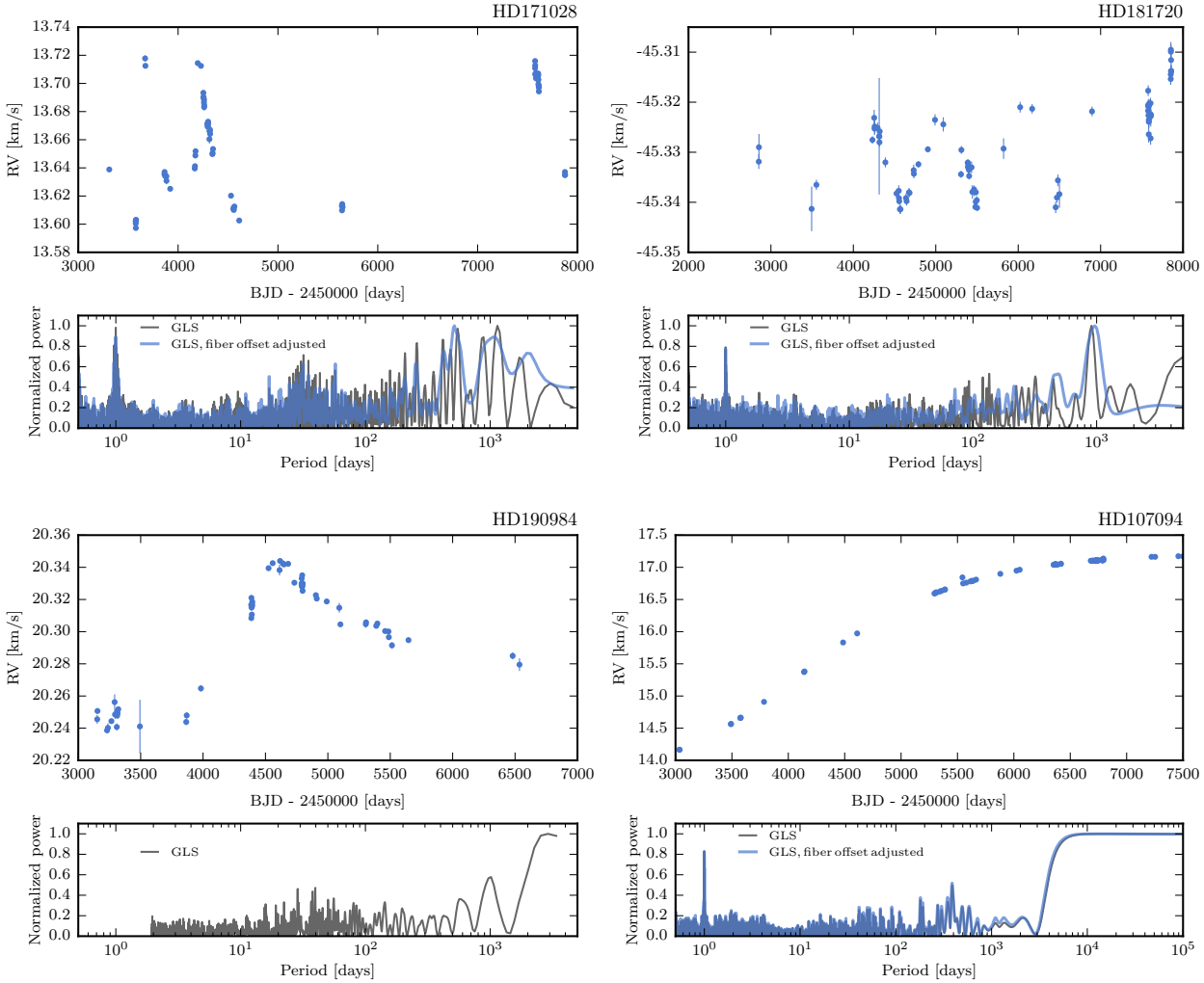


Figure 4.13: RV observations and GLS periodograms for the stars in the LP sample known to host giant planets. When the star has observations after the fiber change, the GLS was calculated taking into account an RV offset.

* ◯ *

The planetary companion to HD 171028 was announced in Santos et al. (2007), with a mass of 1.8 $M_{\rm Jup}$ and an orbital period of 538 days. The discovery paper analysed 19 HARPS RV measurements, obtained within the GTO time, between October 2004 and April 2007. The orbital parameters were revised in Santos et al. (2011), with a few extra RV observations. Since then, we have obtained some more observations, for a total of 72 measurements covering a timespan of almost 12.5 years.

The RV variations of HD 171028 are clear at a glance, and a GLS periodogram identifies a dominant periodicity around 520 days, when the RV offset due to the fibers change is corrected (see Fig. 4.13). The data was analysed with the method presented in Section 2.3. The maximum number of planets was set to $N_{p,\max} = 3$, considering a uniform prior for N_p.

Since the latest (22) observations were made after the HARPS fiber update, we include the parameter δ in the model. An initial analysis with the default prior for δ, between 0 and 30 $\mathrm{m\cdot s^{-1}}$, revealed a posterior distribution clearly skewed towards the upper edge of the prior. The prior range was then increased to 100 $\mathrm{m\cdot s^{-1}}$. We also include a possible long-term trend tr.

The resulting posterior distribution for N_p is shown in Fig. 4.14. with a clear preference for two planets. The posterior distribution for the orbital periods shows a very strong peak around 550 days, but the period of a possible second signal is poorly constrained. To help us pin down the orbital parameters of the known giant planet and the possible second companion, we repeated the analysis fixing the number of Keplerians to $N_p = 2$. A total of 500 000 samples from the DNS target distribution result in 114 151 posterior samples.

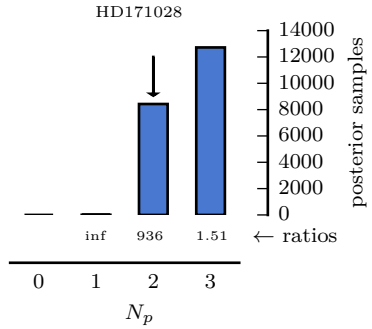


Figure 4.14: Posterior distribution for N_p for the analysis of HD 171028.

Still in this second analysis, the parameters of the second companion are not well constrained. The posterior for the orbital periods shows a peak close to 180 days, which might be caused by the stitching of the HARPS CCD (see the discussion in section 4.3.5). The third highest peak in this posterior (besides the one close to 550 days) is around 70 days, which we attribute to the second companion.

In this situation, we can provide two different estimates for the orbital parameters of the companions. The first and simplest, is to use the maximum likelihood solution, selected from all the posterior samples. This solution gives the orbital parameters shown in Tabel 4.2, but we choose not to provide uncertainties associated with these values, since they are based on one single solution.

The maximum likelihood solution is represented in Fig. 4.15. RV observations before and after the change of fibers are color-coded.

	HD 171028 b	HD 171028 c	
P	554.59	68.76	[days]
K	61.92	5.25	[m·s^{-1}]
e	0.624	0.52	
$m_p \sin i$	2.06 $M_{\rm Jup}$	30.27 $M_\oplus$	

Table 4.2: Some orbital parameters corresponding to the maximum likelihood 2-planet solution for HD 171028.

The middle panel shows the residuals from this fit (note the different units) and the bottom panels in the figure show the phase curves for the two planets, after the other planet was subtracted.

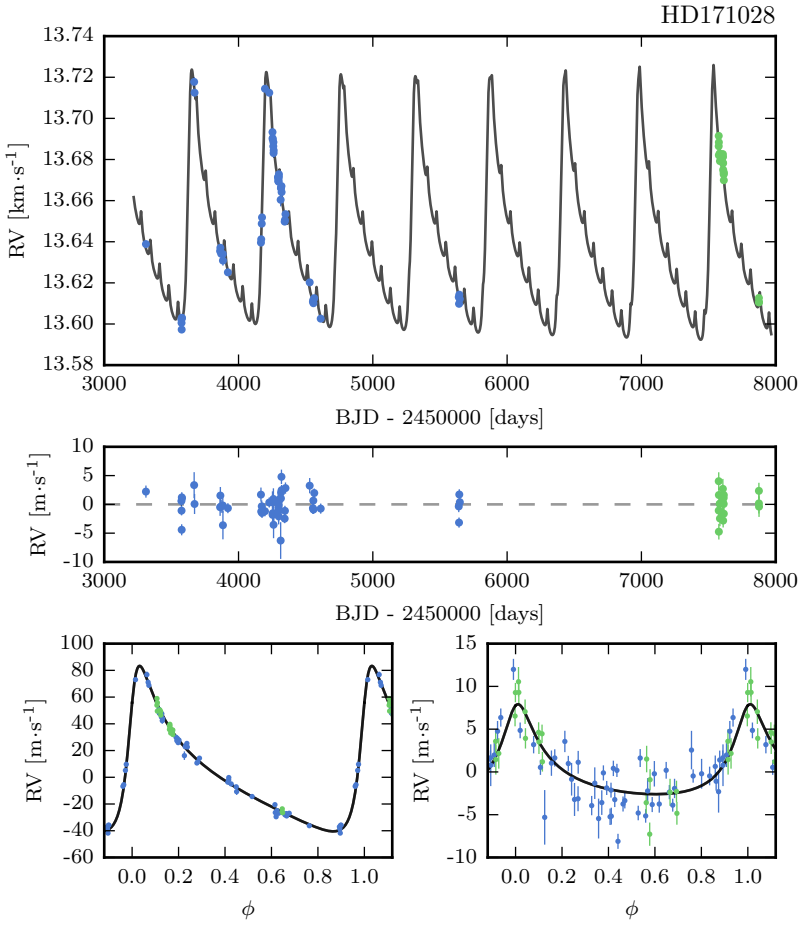


Figure 4.15: The maximum likelihood 2-planet solution for HD 171028. Observations before and after the HARPS fiber change are colored blue and green, respectively. The middle panel shows the residuals from this fit, in m·s^{-1}. In the bottom are the phase curves for the two planet signals, after subtraction of the signal from the other planet.

A second option to obtain estimates for the orbital parameters is to select the samples from within the high posterior regions surrounding the two periods. We do this by applying cuts to the posterior in the orbital period – semi-amplitude space. For the giant planet with the $\sim$550 day period, we expect these to be a correct estimation of the uncertainties, because there is a very clear posterior peak in this region. For the second companion we regard our uncertainties as optimistic, since the posterior does not constrain completely its parameters. This second set of values for the orbital parameters is shown in Table 4.3, where the uncertainties were calculated based on

the 68% percentiles of the posterior samples within the two regions. Note the small differences to the maximum likelihood solution.

	HD 171028 b	HD 171028 c	
P	552.08 ± 1.1	71.70 ± 0.04	[days]
K	$60.73^{+0.97}_{-1.05}$	$4.49^{+1.21}_{-0.88}$	[m·s^{-1}]
e	0.599 ± 0.01	0.55 ± 0.13	
$m_p \sin i$	$2.06 \pm 0.1\ M_{\rm Jup}$	$25.50^{+5.74}_{-4.56}\ M_\oplus$	

Table 4.3: Orbital parameters for the two companions of HD 171028, obtained from the posterior high-probability regions.

Our results suggest very strong evidence for a second planetary companion to HD 171028, besides the already known giant planet. This second planet has around twice the mass of Neptune, and orbits at 0.35 AU from its host.

* ◯ *

Two other giant planets, orbiting HD 181720 and HD 190984, were announced in Santos et al. (2010b) with masses of 0.37 $M_{\rm Jup}$ and 3.1 $M_{\rm Jup}$, moderately eccentric orbits, and long orbital periods.

For HD 181720, a total of 80 measurements were obtained, spanning over 13.5 years (Fig. 4.13). We included both *tr* and δ in the model, since there are 25 RV observations made after the fiber change. The analysis of these data result in the clear detection of only one planet (Fig. 4.16).

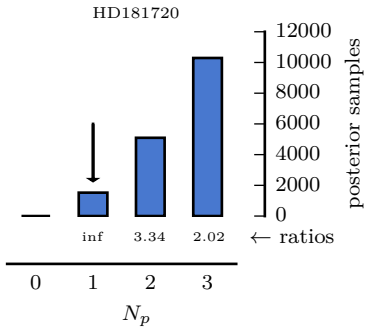


Figure 4.16: Posterior distribution for N_p from the analysis of HD 181720.

The posterior for *tr* includes zero in the 68% credible interval, and therefore the presence of a significant long-term trend is ruled out. For the fiber offset δ, the posterior median and standard deviation provide an estimate of $\delta = 12.46 \pm 2.68\,\mathrm{m\cdot s^{-1}}$. The orbital period and semi-amplitude are larger than in the discovery paper, and the eccentricity is significantly lower (see Table 4.4). Our estimate for the planetary minimum mass of HD 181720 b is of $0.41 \pm 0.03\ M_{\rm Jup}$.

HD 190984, on the other hand, was observed 59 times over 9.2 years. The timespan of the observations does not completely cover one orbital period of the planet, although the eccentric orbit is clear from Fig. 4.13. Repeating the analysis as done previously, with $N_{p,\max} = 3$, results in a significant detection of only one planet.

However, for the analysis of HD 190984, it is not clear a priori if a linear trend should be included in the model and its inclusion changes considerably the resulting solutions. We therefore decided to test two models $A2$ and $A3$, including the parameter *tr* in the model or not, respectively. In these models we used $N_{p,\max} = 1$.

Some of the results are shown in Fig. 4.17, with the posterior predictive samples for the two models. The estimated orbital parameters are presented in Table 4.4. Two very different solutions are obtained if the linear trend is included or not. Unfortunately, the estimated values of the evidence for the two models are similar ($\ln \mathcal{B}_{A2,A3} = 0.3$)

and do not provide a clear preference for $A2$ or $A3$. Therefore, the data currently available cannot constrain the presence of a long-term linear trend.

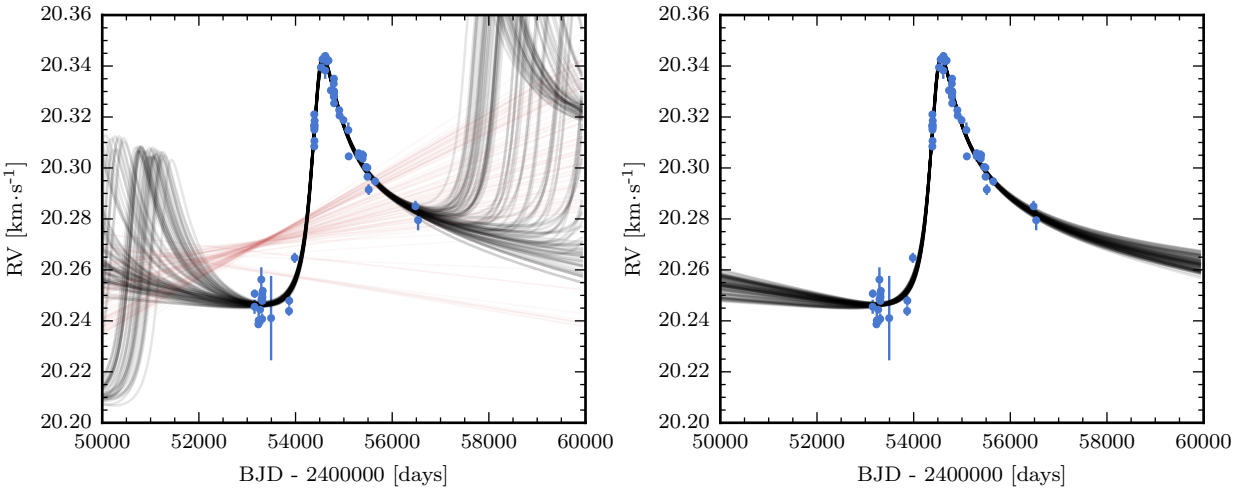


Figure 4.17: Example solutions for HD 190984, with models $A2$ (left) and $A3$ (right). Each panel shows the observed RVs and 100 posterior predictive samples for the Keplerian components (in black) and the associated linear trend (in red), when the latter is included.

If the linear trend is included, we obtain a solution close to the one of the discovery paper (but note that Santos et al. (2010b) did not consider a linear trend in the analysis of this star), even if with a lower orbital period. This solution corresponds to a 2.99 $M_{\rm Jup}$ planet in an eccentric orbit, with the orbital period close to 4100 days. In this model, the slope of the linear trend is well constrained, providing an estimate of $2.84 \pm 1.61\,\mathrm{m\cdot s^{-1} yr^{-1}}$. When the trend is not included in the model, the solution is more eccentric, with an orbital period larger than 27 years (around 10 000 days). Such a planet would have a mass of 3.59 $M_{\rm Jup}$.

* ◯ *

In Santos et al. (2011), a candidate planet for HD 107094 was also announced. Their 14 RV measurements showed a clear long-term trend, though the orbital parameters of a putative planet were difficult to constrain. Their best fit solution corresponded to a 4.5 $M_{\rm Jup}$ planet with $P = 1870$ days, $K = 88\,\mathrm{m\cdot s^{-1}}$, and $e = 0.13$, together with a linear trend with a slope of $402\,\mathrm{m\cdot s^{-1}\cdot yr^{-1}}$.

Another 74 RVs were obtained since then, including 4 measurements after the HARPS fiber change (Fig. 4.13). These new observations allow us to start detecting a curvature on the long period signal. We identify two of the new observations as potential outliers in the RV timeseries at BJD= 55543.87 and BJD= 56793.53 (see Fig. 4.18), and discuss their treatment further below.

For the analysis of HD 107094, we will proceed with a few different assumptions. First, we will analyse the data with $N_{p,\max} = 3$, remov-

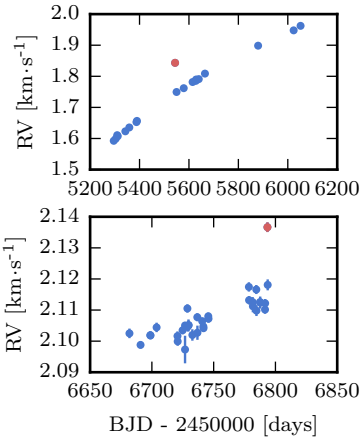


Figure 4.18: Potential outliers in the RVs of HD 107094. Note the different scales of the two panels. The point in the top panel is already noticeable in Fig. 4.13.

ing the two outliers identified before, and considering a linear trend, as in section 4.3.1. We will denote this set of assumptions by $A1$. Afterwards, we will fix $N_{p,\max} = 1$ and test two models, with ($A2$) and without ($A3$) a linear trend. As a final check, we will consider a t-Student likelihood and attempt to include the two outlier points ($A4$) to check if the conclusions remain the same.

The posterior distribution for N_p, with assumptions $A1$, is shown in Fig. 4.19. It points to the significant detection of only one planet, with posterior ratios $(\mathcal{B}_{10}; \mathcal{B}_{21}; \mathcal{B}_{32}) = (\text{inf}; 3.0; 1.5)$. The fit is shown in Fig. 4.20, represented by 100 posterior predictive draws of both the complete signal and the linear trend component.

The posterior samples obtained with this analysis could be used directly to provide constraints on the orbital parameters of the planet and the slope of the linear trend. These estimates will take into account uncertainty in the number of planets N_p, at least within the prior we considered, up to $N_{p,\max}$. In Fig. 4.21, the 1-D and 2-D posterior distributions for the orbital period, semi-amplitude and eccentricity are shown in green, where samples with any number of planets are all combined. Note that the semi-amplitudes and especially the eccentricities are clearly unconstrained.

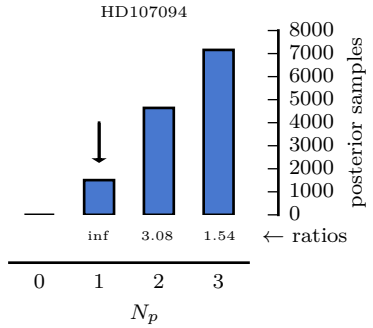


Figure 4.19: Posterior distribution for N_p from the analysis of HD 107094.

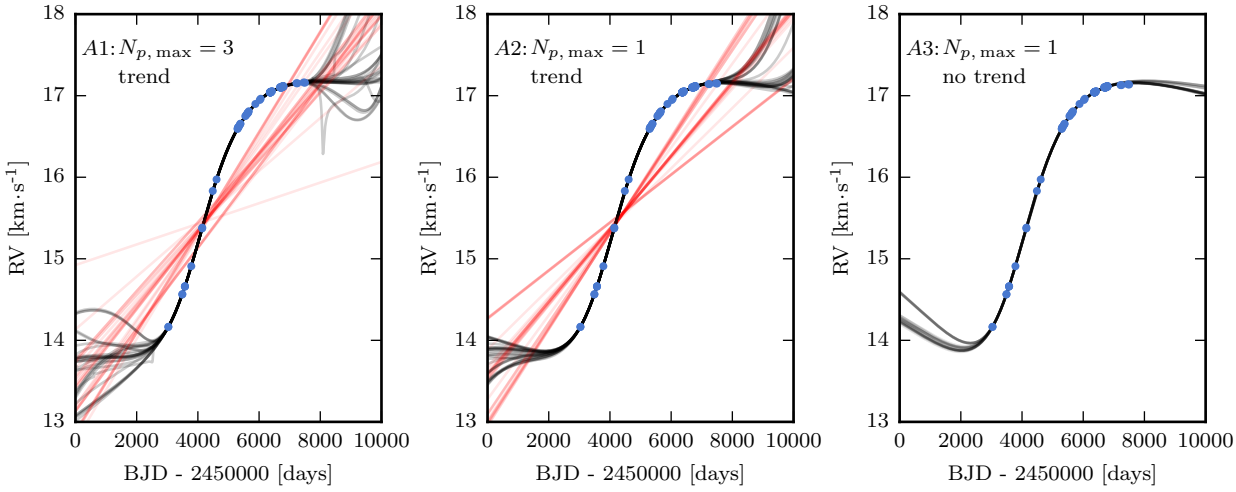


Figure 4.20: Example solutions for HD 107094, with models $A1$ (left), $A2$ (middle) and $A3$ (right). Each panel shows the observed RVs together with 100 posterior predictive samples for the Keplerian components (in black) and for the associated linear trend (in red), when the latter is included.

The data for HD 107094 do not provide, a priori, a clear justification to include or not the presence of a linear long-term trend. Therefore, we can again use the full power of Bayesian model selection, to compare a model with the trend to one without. We consider assumptions $A2$ and $A3$ and obtain posterior samples for each model, calculating the individual evidence values.

The posteriors for N_p are clear in both cases, containing only samples with $N_p = 1$. But the orbital parameters of the detected planet depend strongly on whether or not the linear trend is included. This is clear in Fig. 4.20, which shows the posterior predictive samples

For $A2$ and $A3$, we could have fixed N_p to 1, instead of considering $N_{p,\max} = 1$, but the evidence for one planet is large enough that it does not affect the results.

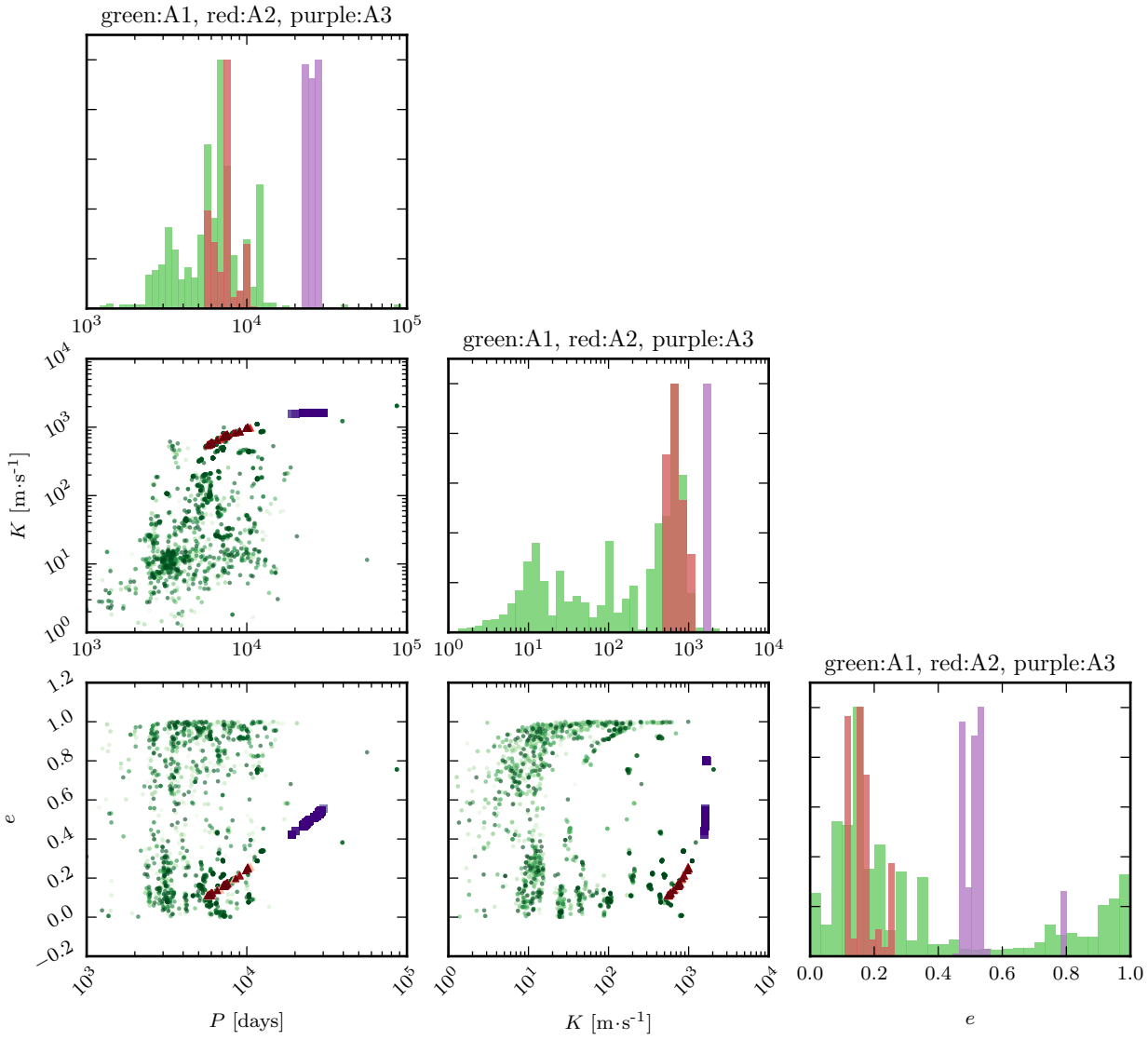


Figure 4.21: Joint and marginal posterior distributions for the orbital period, semi-amplitude and eccentricity from different analysis of the data for HD 107094. Green, red and purple colors correspond respectively to the set of assumptions *A*1, *A*2 and *A*3.

for the two models in the middle and right panel. Without the linear trend, the data is only compatible with solutions that have larger orbital periods and eccentricities. The marginal posteriors for P, K and e are shown in Fig. 4.21 in red and purple, respectively.

For the two models, the evidence values result in a Bayes factor of $\ln \mathcal{B}_{A2,A3} = 3.94$, corresponding to strong evidence in favour of $A2$, the model which includes the linear trend. Nevertheless, in Table 4.4, we report the estimates for the orbital parameters of HD 107094 b, as obtained from both models $A2$ and $A3$.

Regardless of the inclusion of the linear trend in the model, both solutions correspond to brown dwarfs instead of giant planets. With model $A2$ we obtain a companion with 59.12 $M_{\rm Jup}$, with an orbital period close to 20 years and a moderate eccentricity, $e = 0.16$. The slope of the linear trend is estimated as 169.11 ± 32.36, although one

should note here that the posterior distribution for tr is multimodal (see Fig. 4.20, middle panel) and there are strong correlations between tr and some orbital parameters, such as P and e. This means that the uncertainties presented in Table 4.4 might be underestimated. Model $A3$ results in a more massive companion with 180.78 $M_{\rm Jup}$, a period of almost 87 years and a much higher eccentricity.

As mentioned above, we also attempted to model the RVs of HD 107094 using a Student-t likelihood in order to be able to include the two outlier points identified before. In Chapter 2, I mentioned this possibility briefly in section 2.3.5, but did not explore it further. Very briefly, in this model we include an extra parameter ν, representing the degrees of freedom of the Student-t likelihood, and infer its posterior distribution together with the other parameters. For high values of ν, the Student-t distribution approximates the Gaussian distribution while for lower values it allows for more extreme observations.

With assumptions $A4$, a model with $N_{p,\,\rm max} = 3$, and including the two outlier RVs, we recover the same results as before, with only one planet confidently detected in HD 107094's dataset. The estimates for the orbital parameters do not change considerably relative to the other sets of assumptions.

In summary, our analysis of the four targets already known to be giant-planet hosts, confirms the presence of these companions and provides updated orbital parameters. For HD 171028, we find strong evidence for one additional signal, caused by a $\sim$25 $M_\oplus$ companion with a period around 70 days. Our solution for the giant planet HD 171028 b results in a more massive planet at a longer period, when compared with Santos et al. (2007). For HD 181720 b we also find a longer orbital period, a much lower eccentricity, and a slightly higher mass than the solution reported in Santos et al. (2010b). The data for HD 190984 does not constrain a possible linear trend, but its inclusion changes the resulting solutions for the companion. We present two possible solutions for HD 190984 b and wait for new data to be able to choose between them. Finally, for HD 107094 our best solution correponds to a 59.12 $M_{\rm Jup}$ brown dwarf companion with an orbital period close to 20 years.

These results suggest that giant planets in long period orbits may be common around moderately metal-poor stars. Together with the results from the long-term trends, obtained in section 4.3.1, we can place constraints on the population of giant planets in the metal-poor regime. We will come back to this issue in section 4.4.2.

Table 4.4: Orbital parameters for the giant planets detected in the metal-poor sample.

	HD 171028 b	HD 181720	HD 190984 *A*2	HD 190984 *A*3	HD 107094 *A*2	HD 107094 *A*3	
P	553.44 ± 0.26	981.79 ± 8.16	$4102.59^{+2598.95}_{-902.61}$	$10026.56^{+8060.05}_{-2517.01}$	$7285.41^{+1699.40}_{-1334.77}$	$31665.18^{+11123.48}_{-7932.76}$	[days]
K	59.73 ± 1.25	9.06 ± 0.52	44.02 ± 2.59	48.16 ± 1.05	730.10 ± 130.81	1632.89 ± 17.83	[m·s^{-1}]
e	0.59 ± 0.01	0.06 ± 0.04	$0.56^{+0.12}_{-0.08}$	$0.75^{+0.08}_{-0.06}$	$0.16^{+0.05}_{-0.05}$	$0.55^{+0.08}_{-0.06}$	
ω	5.26 ± 0.03	3.02 ± 1.18	5.51 ± 0.07	5.49 ± 0.07	4.32 ± 0.06	4.33 ± 0.06	
T_0	53079.63 ± 1.58	52529.84 ± 219.65	$50442.33^{+815.46}_{-2405.04}$	$44413.65^{+2507.96}_{-8040.69}$	$46524.32^{+1356.11}_{-1570.85}$	$22179.83^{+7791.63}_{-11147.80}$	[BJD]
$v_{\rm sys}$	13.64031 ± 0.00049	-45.33043 ± 0.00038	20.26876 ± 0.00276	20.26860 ± 0.00421	14.91296 ± 0.12283	15.86802 ± 0.05724	[km·s^{-1}]
tr	–	–	2.84 ± 1.61	–	169.11 ± 32.26	–	[m·s^{-1}·yr^{-1}]
δ	23.73 ± 1.11	17.34 ± 1.04	–	–	10.07 ± 5.96	12.35 ± 5.57	[m·s^{-1}]
s	3.04 ± 0.33	1.84 ± 0.22	3.42 ± 0.45	3.41 ± 0.44	2.54 ± 0.32	2.84 ± 0.30	[m·s^{-1}]
$m_p \sin i$	2.05 ± 0.11	0.41 ± 0.03	$2.99^{+0.37}_{-0.25}$	$3.59^{+0.24}_{-0.21}$	$59.12^{+12.53}_{-15.80}$	$180.78^{+10.66}_{-11.78}$	[$M_{\rm Jup}$]
a	1.35 ± 0.03	1.87 ± 0.04	$5.12^{+2.01}_{-0.79}$	$9.35^{+4.43}_{-1.66}$	$6.76^{+0.88}_{-0.85}$	$18.13^{+3.90}_{-2.59}$	[AU]
M_*	1.08 ± 0.08	0.90 ± 0.06	1.08 ± 0.08	1.08 ± 0.08	0.79 ± 0.06	0.79 ± 0.06	[$M_\odot$]
Δt	4565.31	5003.31	3383.75	3383.75	4465.97	4465.97	[days]
N	72	80	59	59	86	86	
$p(D)$	264.22	323.10	210.69	210.99	322.90	318.95	
$\mathcal{H}$	40.31	24.54	30.73	30.51	52.18	50.04	[nats]

4.3.3 HD 175607: *a planet at the stellar rotation period?*

In this section we analyse HD 175607, another interesting target from our metal-poor program. This star was the subject of Mortier et al. (2016), in which we reported evidence for the presence of a small Neptune-mass planet with an orbital period $P = 29.01 \pm 0.02$ days in a slightly eccentric orbit. If the presence of this planet is confirmed, HD 175607 would be the most metal-poor G dwarf with an orbiting Neptune-like planet.

HD 175607 is a bright G6 dwarf at a distance of 45.27 pc (van Leeuwen, 2007). As for the other targets in the metal-poor sample, the parameters of HD 175607 were derived from the sum of individual HARPS spectra, with a total signal-to-noise ratio of 246.4.

The spectroscopic metallicity is -0.62 ± 0.04 dex and the star has a mass of $0.74 \pm 0.05 M_{\odot}$ and a radius of $0.71 \pm 0.03 R_{\odot}$. Some of the spectroscopic parameters are shown in Fig. 4.22, revealing that HD 175607 is a fairly average star within our sample, albeit with low values for the mass and effective temperature. This star probably belongs to the Galactic thin disk, or is transitioning between the thin and thick disks, with a value for [α/Fe] of 0.26 (Adibekyan et al., 2012b). The α-enhancement is a hint that this star is likely to be a planet host, following Adibekyan et al. (2012a).

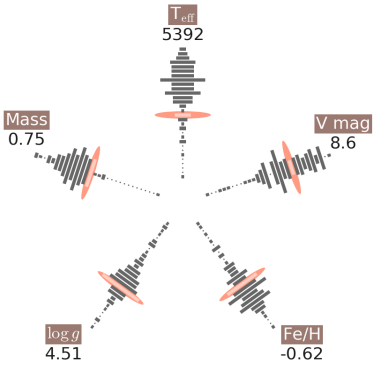


Figure 4.22: Starfish diagram with the spectroscopic parameters of HD 175607.

We can get an estimate for the rotational period with the empirical relationships of Noyes et al. (1984b) or Mamajek and Hillenbrand (2008), via the chromospheric activity indicator $\log R'_{\rm HK}$. The value for the weighted mean of $\log R'_{\rm HK}$ is -4.92 which, when combined with a B-V value of -0.7, gives an estimated rotational period of 28.94 ± 1.02 days or 29.67 ± 7.15 days, from the two calibrations respectively.

A total of 128 HARPS observations are available, adding a few new points since Mortier et al. (2016) was published. These observations were made on 115 individual nights, between July 2004 and April 2015. The RVs derived using the HARPS DRS and the GLS periodogram are shown in page 174 in Appendix C (see also Fig. 4.23).

Mortier et al. (2016) used different methodologies for the study of RV time series. First, the correlations between the RVs and the activity indicators were studied, looking for hints of activity-induced variations that could mimic a planetary signal (since there is a strong peak in the periodogram of the RVs very close to the estimated rotation period). No strong correlations were found (see Fig. 4.23) and the indicators do not show significant periodic variations.

Both a pre-whitening analysis of the RV periodogram and an MCMC algorithm suggested a likely 2-planet model, with $P_b \sim 29$ days and $P_c \sim 1300$ days. The RVs were also analysed with a Gaussian process model, identical to the one described in Rajpaul et al. (2015), resulting in a significant detection of only one planet at $P_b \sim 29$ days, but with a lower mass.

With the methodology described in Section 3.3, I analyse here the complete dataset for HD 175607, in order to confirm or refute the detection of the announced planet. The GP model for stellar activity

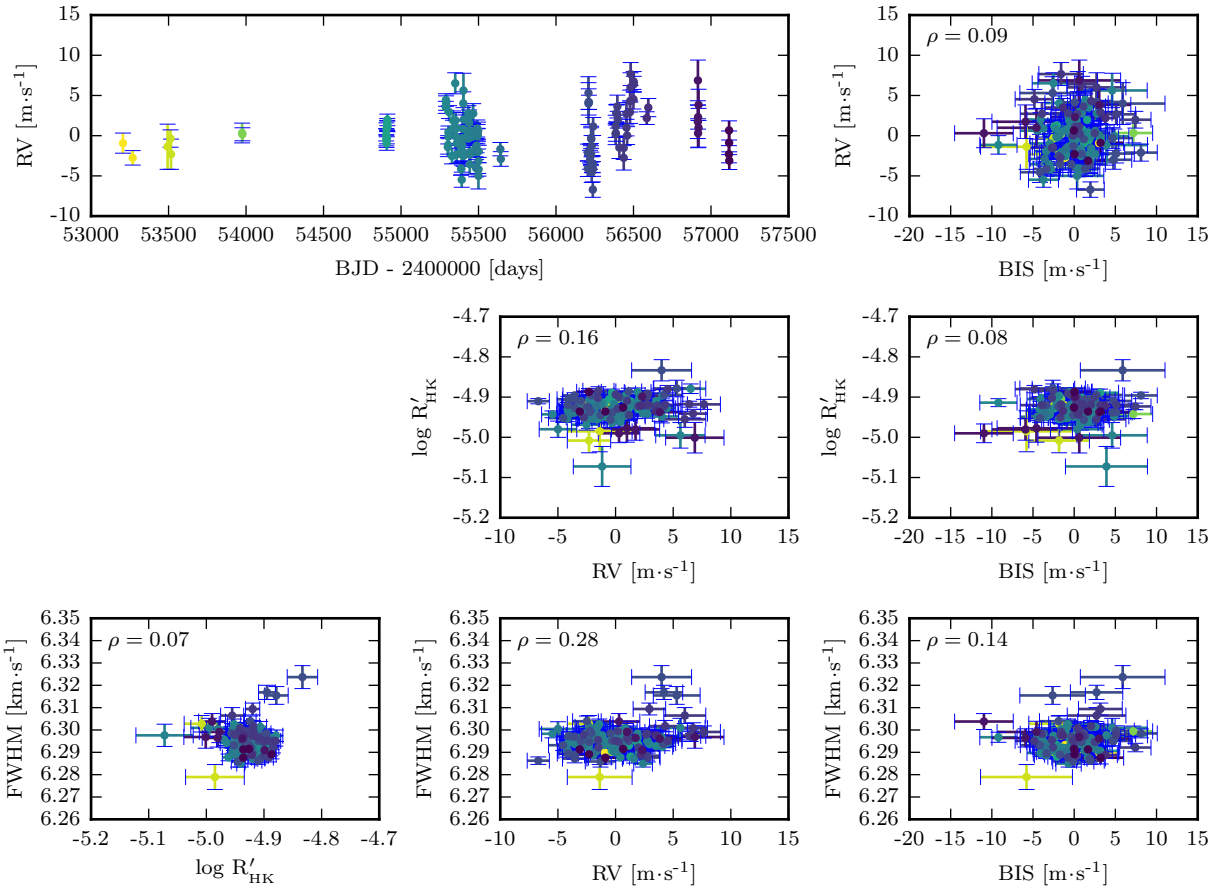


Figure 4.23: RV time series and correlations between the RVs and some activity indicators for HD 175607. Each correlation panel presents the value of Pearson's correlation coefficient ρ.

variations is included, and the maximum number of planets was set to $N_{p,\max} = 5$. A total of 93622 samples from the joint posterior distribution were calculated using the DNS algorithm.

The posterior for the number of planets N_p is shown in Fig. 4.24. The results point to no planets being detected in these data, as the ratio of probabilities between $N_p = 0$ and $N_p = 1$ is only 1.25. This is at odds with the conclusion of Mortier et al. (2016) and therefore deserves an in-depth discussion.

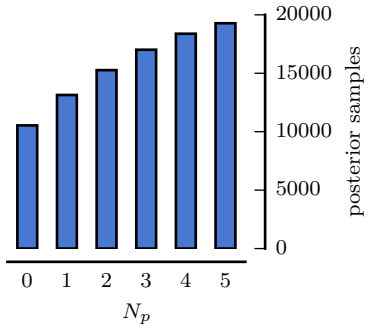


Figure 4.24: Posterior distribution for the number of planets, for HD 175607.

The posterior distributions for a few parameters, both from the Keplerians and the GP, are shown in Fig. 4.25. In these distributions, all the samples with different values of N_p are included, such that we see clearly the prior distributions. Note how the posterior for the orbital period has a strong peak around 30 days, close to the period of the reported planet (and also to the estimated rotation period). We also find that the posterior for η_3 is strongly bimodal, with peaks near 17 days and 30 days.

Note also that, even if their posteriors are mostly unconstrained, the MAP values for the orbital period, semi-amplitude and eccentricity are similar to the values reported for the planet in Mortier et al. (2016). This means that our analysis detects the same signal (which was attributed to the planet by those authors), except that we find it to be

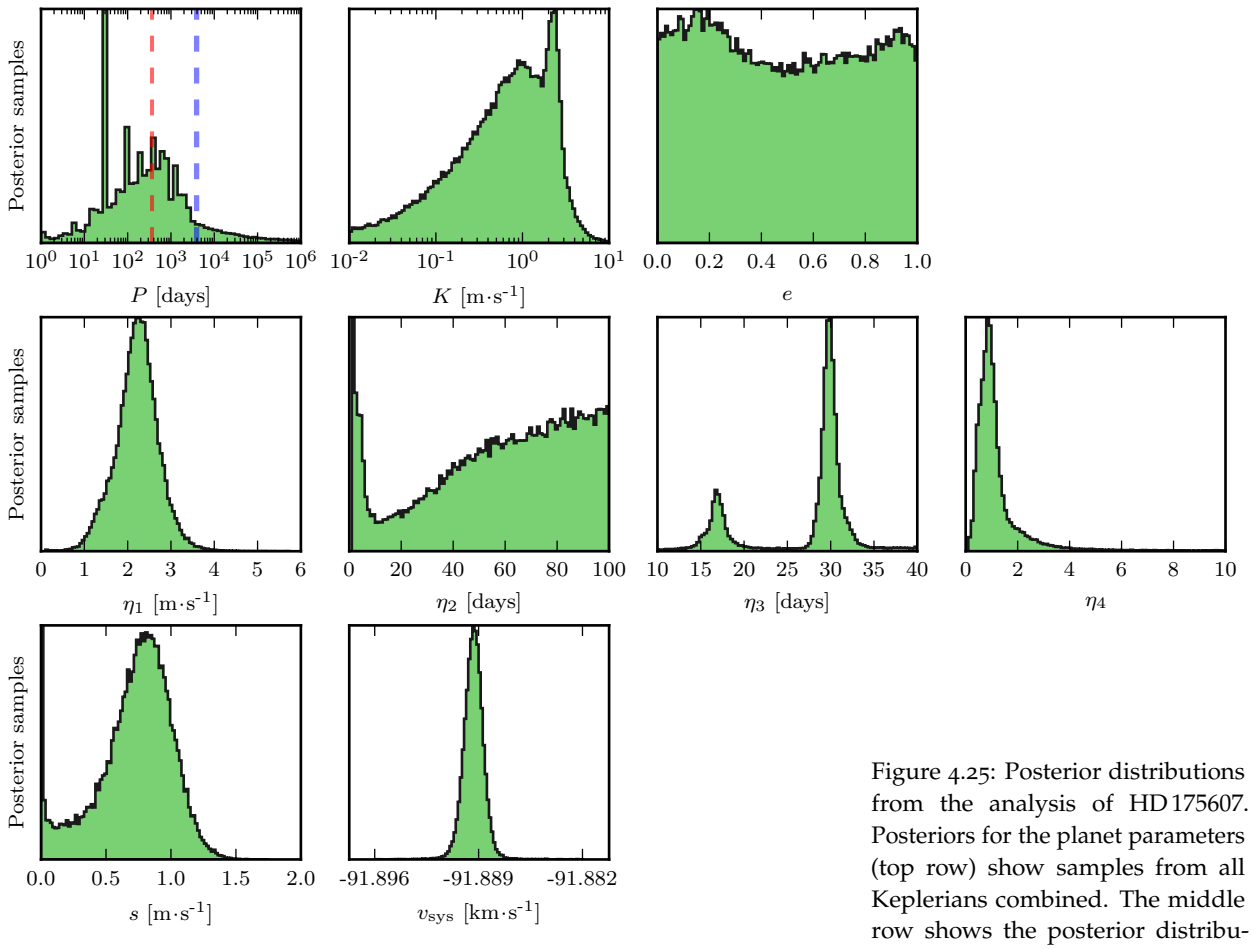


Figure 4.25: Posterior distributions from the analysis of HD 175607. Posteriors for the planet parameters (top row) show samples from all Keplerians combined. The middle row shows the posterior distributions for the hyperparameters and the bottom row for the jitter and systemic RV.

less significant and not a confident detection.

It is also interesting to look at the joint distribution of P_1 (the period of the first Keplerian signal) and η_3, shown in Fig. 4.26. When the orbital period of the planet is near $P \sim 30$ days, the periodic timescale of the GP is much more likely to get the smaller value of 17 days. This means that, within our model, the scenario of having *both* the planet orbital period and the rotation period very close to 30 days is disfavoured. But note also that 17 days is close to the first harmonic of the rotation period, which suggests that the solutions where the Keplerian is at 30 days do not completely model the power at the first harmonic.

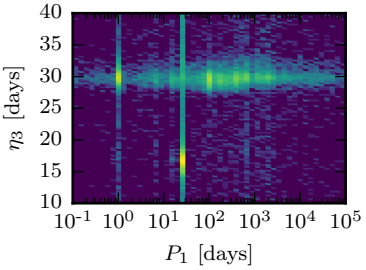


Figure 4.26: Joint posterior distribution for P_1 and η_3.

These results seem to contradict those obtained in Mortier et al. (2016) with the GP analysis, despite, at first, the two models appearing quite similar. There are a few obvious differences though: here we analyse 9 extra RV observations, and we do not consider the information in the activity indicators. We also use different priors and perform the model comparison in a different way.

In order to try to explore these last points, I first selected from all the MCMC samples, those containing either 0, 1 or 2 planets and, of those, the ones with the maximum value of the likelihood. With these maximum likelihood values, the AIC and BIC were calculated. The results of AIC favour the 3-planet model and the BIC favours the 1-planet model, both with significance (i.e. ΔAIC and ΔBIC higher than 10). Thus, the analysis of HD 175607 is another instance where the model preferred by the AIC and BIC disagrees with the one favoured by our method, in addition to the two information criteria disagreeing between each other. In Mortier et al. (2016), the AIC was used as a criteria for model comparison, which can probably explain in part the differences we observe.

Another very important point in the analysis of HD 175607, is the prior for the orbital period (of the proposed planet). In Mortier et al. (2016) we used a very narrow prior, between 27 and 32 days. To what extent does this influence the results? In order to test the influence of this prior specifically, I consider the same model as before except for the use of a uniform prior for the orbital period(s), between 27 and 32 days. To be clear, nothing else was changed in the model apart from the orbital period prior.

After obtaining 102 428 posterior samples with DNS, the resulting posterior distribution for N_p is shown in Fig. 4.27, which can be compared directly to that of Fig. 4.24. The ratio between the probabilities of $N_p = 0$ and $N_p = 1$ is now higher than 30, corresponding to strong evidence for the 1-planet model (although it still does not meet our criteria for a confident detection; Table 2.2). The remaining probability ratios almost did not change.

This shows the nasty effect of considering a prior that is too narrow. The evidence is an integral over the complete parameter space of the model. This parameter space for the 0-planet model, is not altered when changing the prior for the orbital periods. But for the 1-planet model, the volume of the parameter space is much

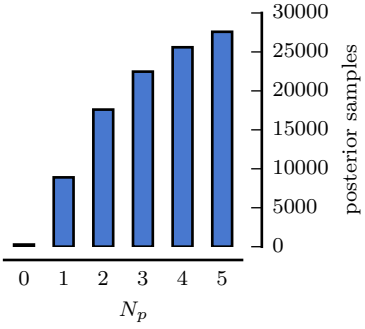


Figure 4.27: Posterior distribution for the number of planets, for HD 175607, with a uniform prior $\mathcal{U}(27, 32)$ for the orbital periods.

smaller with the narrow prior. But the goodness-of-fit of both models (their maximum likelihoods) did not change (considerably) with the changing prior. Therefore, the evidence for the 1-planet model is "artificially" enhanced by the narrow prior.

The results of this analysis do not allow for a confident detection of the reported HD 175607 b. In the model considered here, the scenario where the orbital period of the planet is essentially the same as the rotation period is strongly disfavoured. The RV data nevertheless show a periodicity around 30 days, which is not present in any of the activity indicators. But we do not find this signal to be significant enough to pass our detection criterion.

We also saw how the narrow prior considered by Mortier et al. (2016) can affect the results and increase the evidence ratio between the 0-planet and the 1-planet model. This serves as a cautionary tale for future analyses. It goes without saying that these results are only valid to the extent that our GP component can reproduce the RV variations caused by stellar activity.

4.3.4 HD 41248*: mean motion resonance or stellar activity?*

This section is devoted to the analysis of the HARPS data for HD 41248, the subject of a recent debate over the existence of two orbiting planets. Jenkins et al. (2013) used the first 62 public RV measurements of HD 41248 to announce the detection of a system of two Neptune-mass planets with orbital periods of 18.36 and 25.65 days. These periods place the planets close to a 7:5 mean motion resonance, with implications to their formation mechanism.

With over three times more data, Santos et al. (2014) found no evidence for the planetary system. One of the signals ($\sim$ 25 days) was shown to be related to the rotational period of the star, since it was clearly seen in some of the activity indicators (see Fig. 4.28). The other planet was not consistently found in the new dataset. However, a re-analysis by Jenkins and Tuomi (2014) again found the two proposed planets, with almost exactly the same orbital parameters, and using the complete RV dataset.

HD 41248 is a $V = 8.82$ magnitude G2 dwarf in the constellation Pictor, at a distance of 52 ± 2 parsec from the Sun (van Leeuwen, 2007). The stellar parameters for this star were derived as before (Sousa et al., 2011a), resulting in a temperature $T_{\rm eff} = 5713 \pm 21$ K, surface gravity $\log g = 4.49 \pm 0.05$ dex, and a spectroscopic metallicity $[{\rm Fe/H}] = -0.37^{\pm 0.05}_{\pm 0.01}$ dex.

* ◯ *

The complete dataset for HD 41248 contains 228 RV measurements, spread over a timespan of 10.2 years. The timeseries is shown in Fig. 4.28, together with the correlations between the RVs and the activity indicators. The strongest correlations are found between the RVs and the $\log R'_{\rm HK}$ (with $\rho = 0.29$) and FWHM of the CCF (with $\rho = 0.33$). The periodograms of these three quantities (Fig. 4.29) show

the same significant periodicity around 25 days. This was one of the arguments used in Santos et al. (2014) to show that this periodicity might be related to stellar activity.

From the activity-rotation relations, and with a weighted mean of $\log R'_{\rm HK}$ equal to -4.89, we estimate the rotation period to be 18.16 ± 1.02 days (Noyes et al., 1984b) or 18.36 ± 3.30 days (Mamajek and Hillenbrand, 2008). Note how these estimates are very close to the first planet's orbital period from Jenkins et al. (2013).

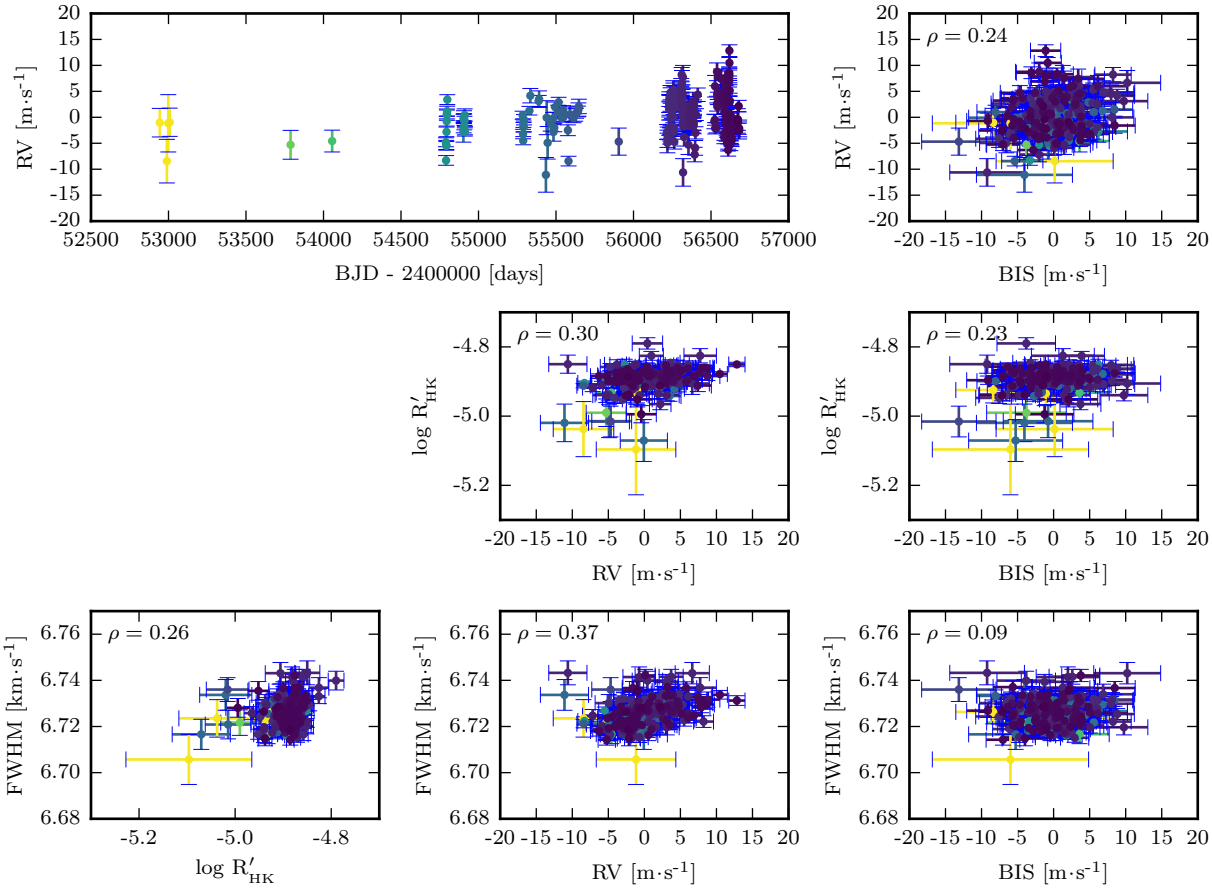


Figure 4.28: RV time series and correlations between the RVs and some activity indicators for HD 41248.

These data present another opportunity to run our model of planet detection. With the same methodology as before (Section 3.3), I analysed the complete dataset of HD 41248. The maximum number of planets was again set to $N_{p,\max} = 5$, with a uniform prior for N_p. The DNS algorithm ran until 10^6 samples of the target distribution were obtained, which correspond to 197 714 samples from the joint posterior distribution. The resulting posterior distribution for N_p is shown in Fig. 4.30, and points to no planets being detected in this dataset: $N_p = 0$ is the mode of the distribution.

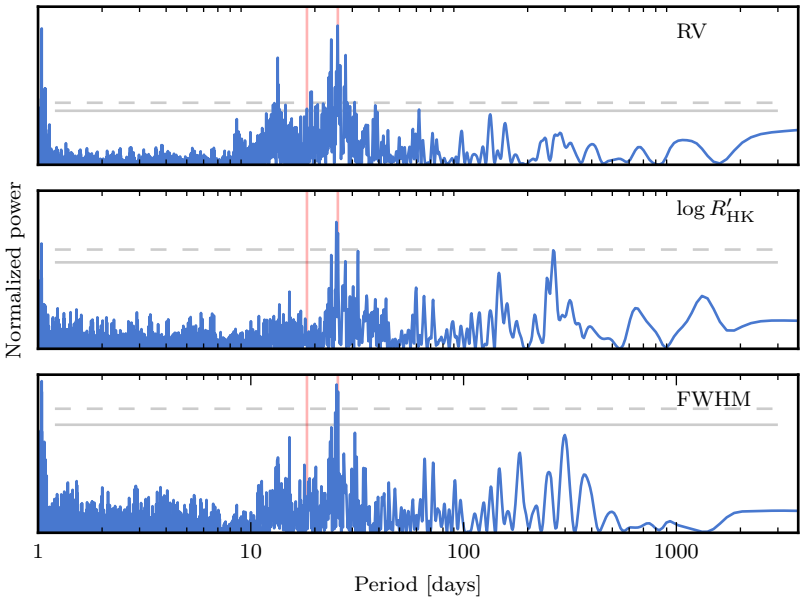


Figure 4.29: GLS periodograms of the RVs, log R'_{HK} and FWHM time series of HD 41248. The orbital periods of the two planets proposed by Jenkins and Tuomi (2014) are marked with red lines. The levels of 1% and 0.1% FAP are shown in solid and dashed gray lines, respectively.

It is interesting to note that the MCMC explored well the parameter space between $P \sim 10$ and $P \sim 40$ days, and that a number of signals are found in this range (Fig. 4.31). The signals at 13 days and 25 days are the strongest, followed by a peak at 18 days and one other close to 30 days, which were all already discussed in Santos et al. (2014). Fig. 4.31 shows also the periods and semi-amplitudes of the two planets suggested by Jenkins and Tuomi (2014), showing that we also find these signals in the data.

The posterior distributions for the GP parameters are very well constrained (see Figs. 4.31 and 4.32). The distribution for η_3 provides an estimate of the rotation period of 25.52 ± 1.13 days, calculated as the median and standard deviations of the posterior samples. Note also (in Fig. 4.31 for example) that the parameter η_1 is constrained around the same value as the amplitude of the proposed planet. This shows that the RV variations with periods of ~ 25 days are being completely explained by the GP, without the need for the Keplerian signal. And, perhaps more importantly, also without the need for another Keplerian at 18 days.

Our results give rise to various questions. From the activity-rotation relations, we have estimated HD 41248's rotation period to be close to 18 days. But the periods found in the log R'_{HK} and in the FWHM are close to 25 days and the GP analysis overwhelmingly prefers this latter value as the activity-related period. These are, coincidentally, the two periods of the proposed planets.

It is interesting to speculate, as Santos et al. (2014) did, on whether differential rotation could cause the rotation period of this star to change by $\sim 30\%$, or if different groups of (evolving) active regions at different latitudes could induce the two periodicities seen in the RVs. We have not answered these questions here, as we could only

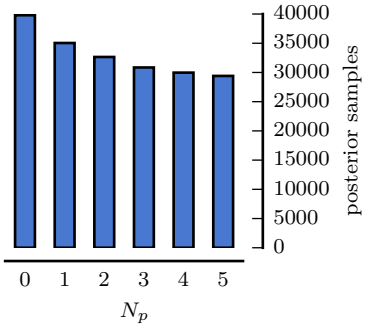


Figure 4.30: Posterior distribution for the number of planets, for HD 41248.

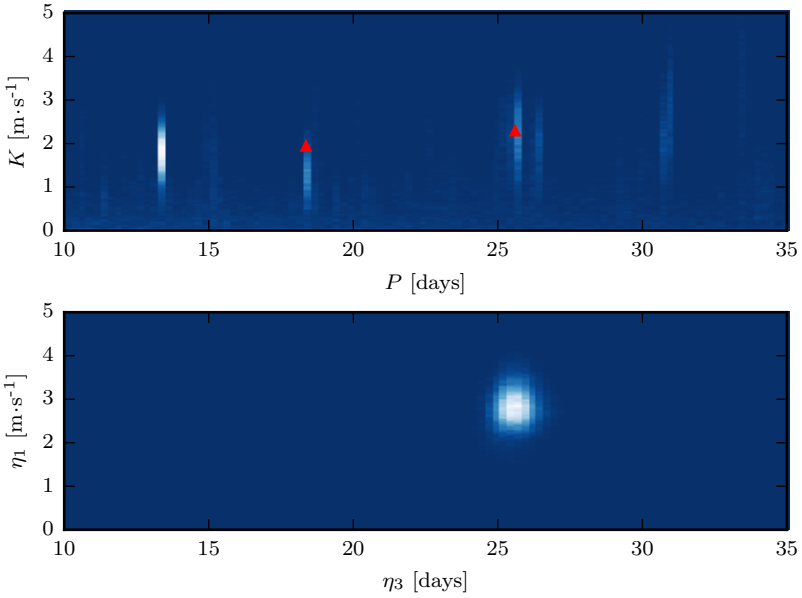


Figure 4.31: Posterior distributions from the analysis of HD 41248. The semi-amplitude and orbital period of all Keplerians is shown in the top panel, and η_1 and η_3 in the bottom one. Parameters of the two planets proposed by Jenkins and Tuomi (2014) are also marked.

do that with a complete model of the distribution of active regions on the surface of the star. Such an analysis would also have to include prescriptions for active region evolution and differential rotation. A very interesting step in this direction was presented recently by Giguere et al. (2016), who used simultaneous RV and photometric measurements to infer the spot distribution of the active K dwarf ϵ Eridani. We plan to develop a similar model and apply it to the observations of HD 41248.

But even from the data analysis point of view alone, it is interesting to wonder why Jenkins and Tuomi (2014) still detect the same two significant planetary signals in the complete dataset. The main differences between our analysis and theirs are (*i*) the models used for correlated noise – Gaussian processes *versus* moving average – and (*ii*) the fact that Jenkins and Tuomi (2014) considered linear correlations with some activity indices directly in their RV model.[5] Even if these approaches have proven to be successful in simulated datasets (see Dumusque et al., 2017), their correctness in real cases is still debated. In particular, a very recent analysis by Feng et al. (2017) used identical models and the same dataset as Jenkins and Tuomi (2014), but reached different conclusions on the detected planets.

HD 41248 provides a good example of the difficulties inherent to the analysis of RV data, especially when searching for very low mass planets. The amplitudes of the RV signals caused by stellar activity are, even in this relatively inactive star, of the same order of magnitude as the expected signals from those planets.

I may note here that a new generation of near-IR spectrographs is currently being developed to provide complementary spectroscopic measurements using other wavelengths, which will certainly help in the characterisation of planetary signals (e.g. Huélamo et al., 2008; Prato et al., 2008; Figueira et al., 2010).

[5] See Bonfils et al. (2007) for a case where these correlations are clearly not linear.

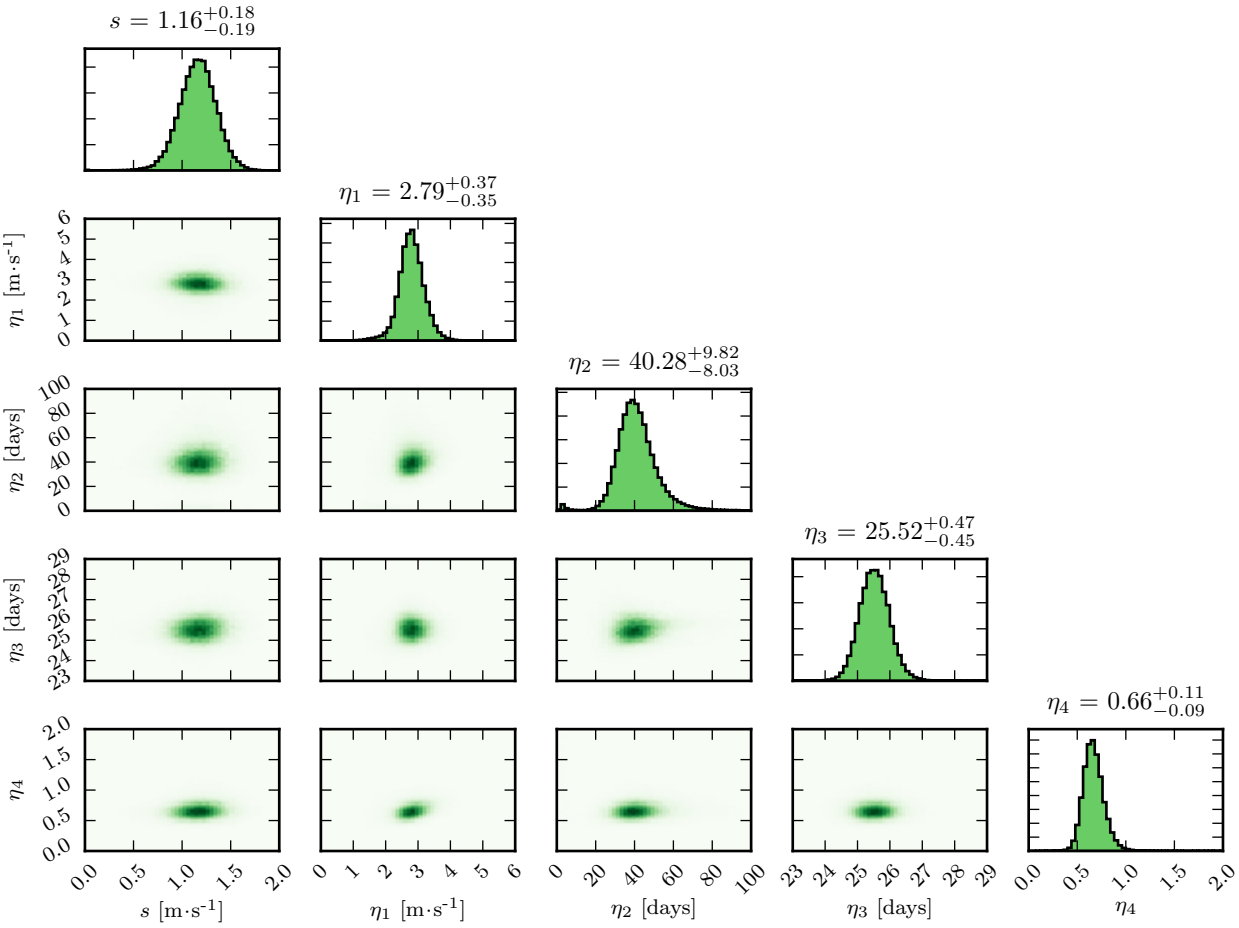


Figure 4.32: Marginalized and joint posterior distributions for the GP parameters, from the analysis of HD 41248. The value at the top of each distribution represents the median and 16% and 84% quantiles of the posterior samples.

4.3.5 *Spurious one-year signals*

While being one of the most precise spectrographs currently in operation, HARPS is most definitely not free of instrumental effects. One such effect was described by Dumusque et al. (2015b). The RVs of several stars observed with HARPS over a period of several years, present a low-amplitude signal at a period of one year (365 days), or its first harmonics ($\sim$180 and $\sim$120 days).

These signals were found to be caused by a discontinuity in the wavelength calibration. The HARPS CCD detector is composed of 32 blocks of 512×1024 pixels, but the pixels between each block do not have the same size as those inside the blocks.[6] When stellar lines cross the block stitchings due to the motion of the Earth around the Sun, this can introduce a RV signal. Usually, this signal has a period of one year or six months, even though other periodicities are also possible, depending on the way in which the lines cross the CCD stitchings (see Dumusque et al., 2015b, for details).

[6] These discontinuities are created by the way in which CCDs are built, with different blocks of pixels being put together by a mechanical claw that deposits them into a common structure.

In the LP sample, we detect this type of signal in a number of RV time series. In some cases, the GLS periodogram of the RVs shows a peak close to one year or its harmonics, and in other cases, an analysis with our RV model finds evidence for a "planetary" signal and the posterior for the orbital periods shows clear peaks at or close to 365, 180 or 120 days.

Dumusque et al. (2015b) proposed two alternative methods to suppress this type of signals: (*i*) either recalculate each CCF and RV with a new correlation mask which does not include the lines that cross the CCD stitchings or (*ii*) fit a sinusoidal signal with a period of 365.25 days to the RV data[7] The first method is clearly preferred, since it is expected to completely correct for the underlying effect. However, it relies on repeated runs of the HARPS reduction pipeline and, more importantly, is currently only possible with substantial help from the HARPS collaboration.

[7] A third option, also mentioned in Dumusque et al. (2015b), would be to include the CCD pixel sizes at all the block stitchings as a free parameter during the wavelength calibration, and derive a corrected wavelength solution. However, only a ghost citation was provided, which I will not repeat.

As a preliminary test we tried to use the second approach by fitting and removing a sinusoidal signal with a period of 365.25 days from the timeseries of the stars in which this effect is detected. The amplitude and phase of this signal is free, as is a RV offset to adjust the mean of the signal and, for some stars, an offset for the observations after the HARPS fibers change. Only the period is fixed.

Our tests show that this approach is only marginally successful. For many targets, the posteriors for the orbital periods still show clear peaks at 365, 180 or 120 days, even after this initial correction is applied. We also attempted to fit and remove a sinusoidal signal with a period of $365.25/2 = 182.625$ days instead since, in some cases, this is the dominant peak found in the posteriors. Even if slightly better results were obtained for a couple of stars, we are not able to draw clear conclusions from these tests.

Therefore, we decided to bite the bullet and re-derive the RVs of a few targets using a custom correlation mask. To build these masks, we first identify the wavelength that correponds to each CCD stitching,

at zero barycentric Earth radial-velocity (BERV). We then removed all lines that are within 10 $km \cdot s^{-1}$ of the minimum and maximum BERV. The 10 $km \cdot s^{-1}$ margins help in also removing lines that are just partially affected by the stitching. After building the custom masks with fewer lines, the HARPS pipeline is used to derive the RVs for a given target (see section 2.2). New RVs were obtained with this procedure for HD 111777, HD 196877, HD 133633, HD 199847, HD 88725, HD 207869, HD 199288, HD 22879, and HD 56274.

4.3.6 *Ensemble analysis*

In the previous sections we have discussed some interesting targets, and presented a few extensions to the RV model presented in Chapters 2 and 3. As a last step before we can search for planets in the complete metal-poor sample, I will introduce two new criteria for the detection of Keplerian signals. These criteria follow closely those reported by Tuomi et al. (2014b).

1. **Posterior ratio criterion** As already mentioned in section 2.3.8, we require that the probability of N_p should be at least 150 times greater than the probability of $N_p - 1$.

2. **Constrained orbital period** The posterior distribution for the orbital period of a candidate signal must be well constrained and unimodal.

3. **Non-vanishing amplitude** The posterior distribution for the RV semi-amplitude of a candidate signal should not include zero inside the 99% credible interval.

Before we justify these new criteria, it is already clear that the first and third criteria will be more easily assessed quantitatively than the second one. It is also clear that the three criteria are related, and should complement each other. A signal that passes all three detection criteria, can be considered to be robustly detected and, very likely, a planet.

The fact that we require the orbital period posterior to be constrained is easy to justify. If this posterior is clearly multimodal, it means that two or more solutions can explain the data, sometimes equally well. Even if not equally well, choosing one of those solutions to claim a detection would underestimate our uncertainty on the orbital parameters. Also, if the orbital period posterior is unimodal but the (relative) uncertainty on P is large, we cannot claim that those RV variations are indeed of periodic nature.

Requiring a non-vanishing amplitude for a detected signal, further safeguards against an imperfect modelling of quasiperiodic noise features present in the data.

In practice, we enforce the second criterion by calculating the 68% percentile range of the posterior distribution for P, and comparing it to the posterior median. We assess multimodality by visual inspection of the posterior distribution. A non-vanishing amplitude is checked

by calculating the 99% percentile ranges of the posterior distribution for K. As we will see, the analysis of our RV timeseries shows that the constrained period criterion is the hardest to satisfy.

* ◯ *

We will use the model developed in Chapter 2 (that is, without the activity component) for the analysis of all 109 stars in our sample, considering a number of default settings. First, we use the reference priors of Table 2.1, considering always $N_{p,\max} = 3$. For those stars which have observations after the HARPS fibers change, we include the parameter δ in the model, and for those in which a linear trend is detected, we include the parameter tr (see section 4.3.1). For some stars we analyse RV data re-derived with a custom mask that avoids the HARPS CCD stitching (see section 4.3.5).

We obtained 50 000 samples from the DNS target distribution for each of the stars, letting the code determine automatically the number of levels required. Whether or not the algorithm "converged" was determined heuristically, as explained in Appendix A. Note that a fixed number of samples from the target distribution, means that we obtain different number of samples from the posterior distribution for each star. As a rule of thumb, the number of samples is much larger than necessary for stars with few observations, but large enough to characterize the posteriors for stars with the largest number of RV observations.

For the sake of consistency, the stars that were already mentioned in previous sections (HD 171028, HD 181720, HD 190984, HD 107094, HD 11397, HD 144589, HD 197083, HD 93351, HD 175607 and also HD 41248) are still included in the ensemble analysis below, but I will not discuss them any further. Nevertheless, for these stars it is still an interesting exercise to check how the posterior distributions for N_p behave within the more simplistic model, which does not include the stellar activity component, for example.

In Fig. 4.33 and 4.34, all the 109 resulting posteriors for N_p are represented, with the posteriors normalized to their maximum values. All stars that have been discussed before are presented with a purple box. For the stars that pass the posterior ratio criterion, the distributions are coloured in purple. In each of these, an arrow marks the number of signals detected. Note that some posteriors show evidence for the presence of at least two Keplerian signals.

In both figures, the stars are sorted by the number of observations from top to bottom and left to right. Stars in Fig. 4.33 have between 5 and 50 observations and in Fig. 4.34 between 51 and 264 observations. One clear result (though maybe not surprising) is that the number of purple distributions, showing the probable detection of planetary signals, is higher in Fig. 4.34.

Apart from the stars already discussed in previous sections, our analysis finds evidence for possible planetary signals in a number of other targets, namely HD 71685, HD 197197, HD 126803, HD 206998,

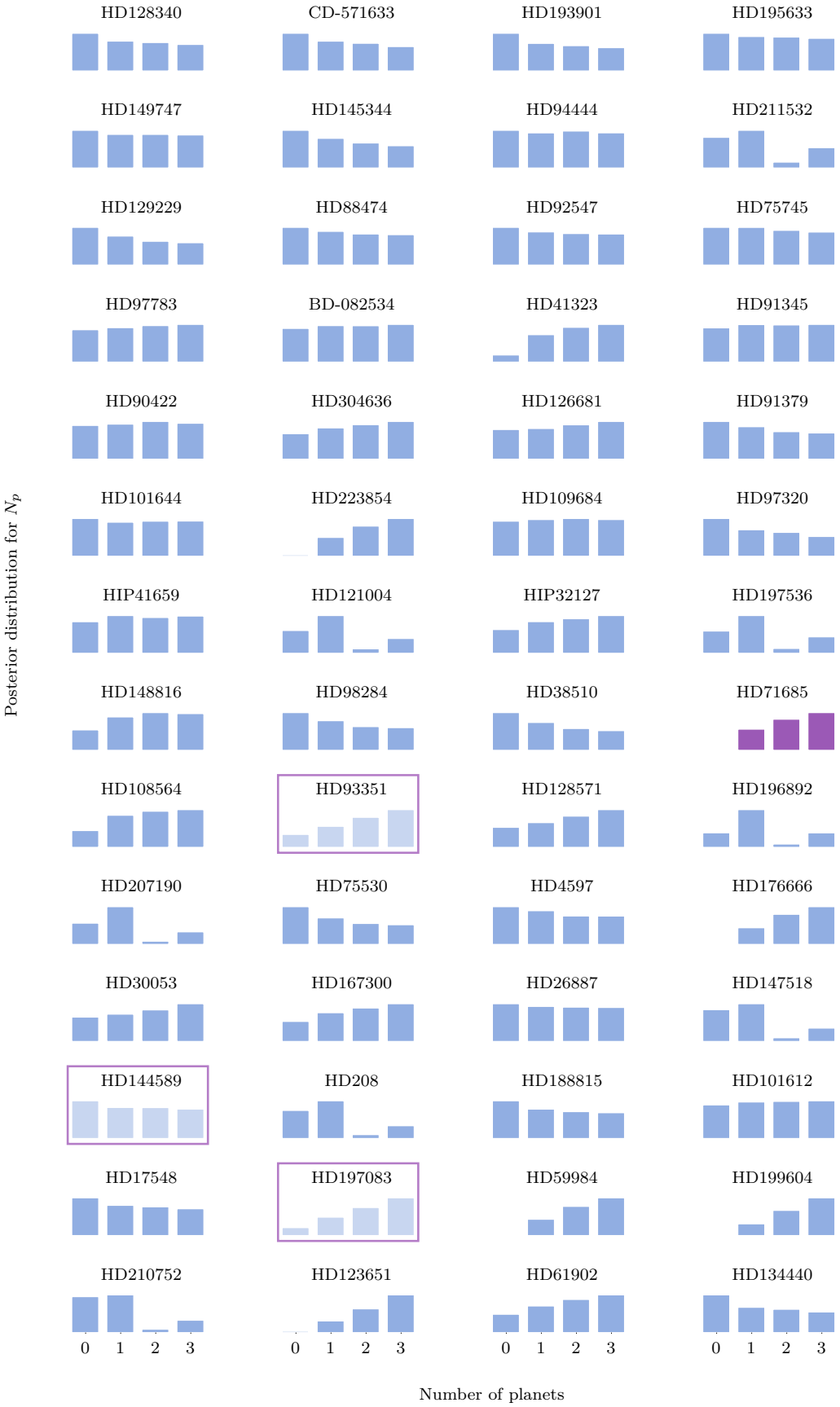


Figure 4.33: Posterior distributions for N_p. The stars are sorted by number of measurements (left to right and top to bottom). Stars analysed in the previous sections are marked with a box and when $\mathcal{B}_{1,0} > 150$, the distribution is shown in purple.

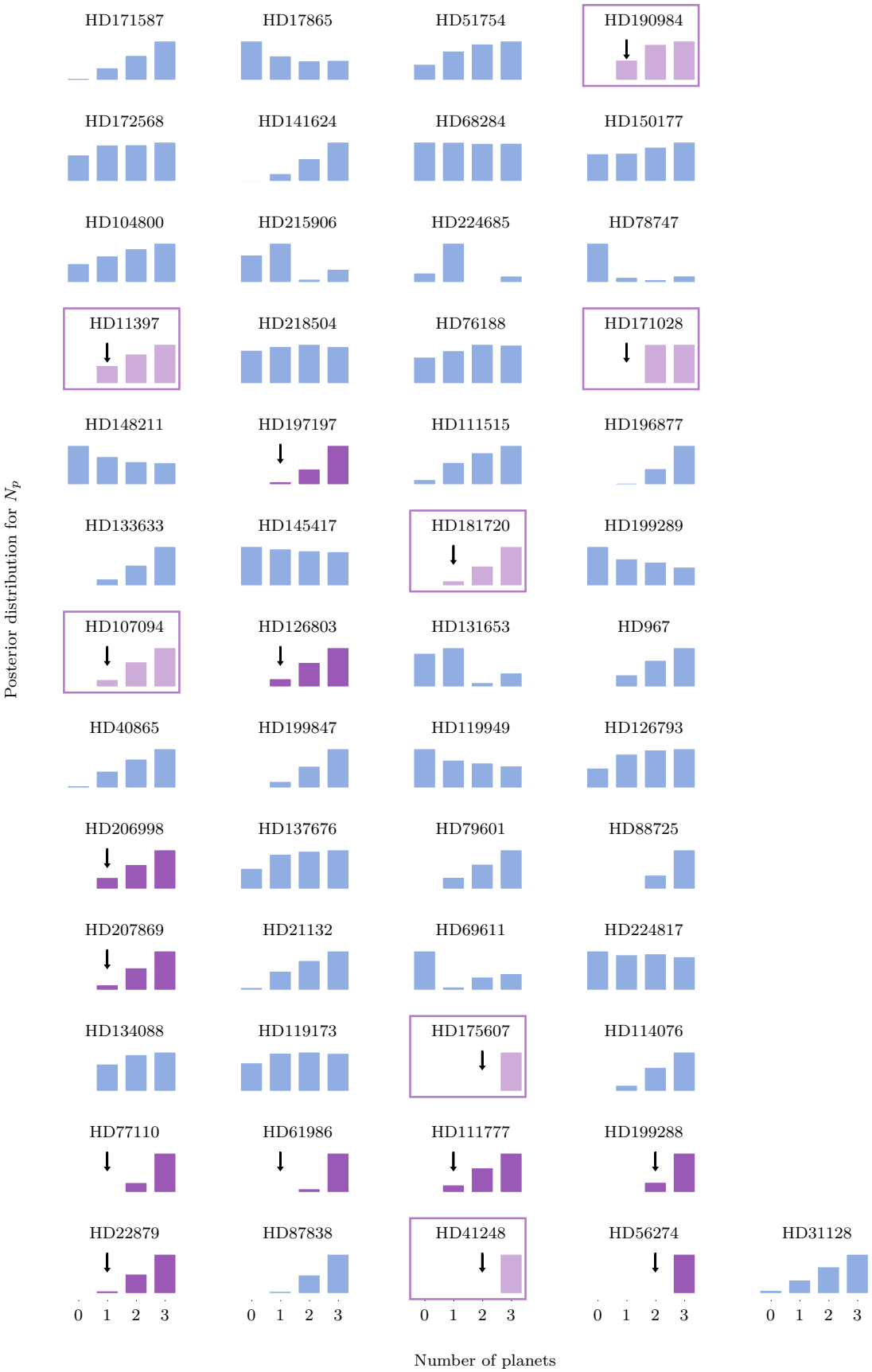


Figure 4.34: Posterior distributions for the number of planets N_p, as in Fig. 4.33 but for the remaining stars.

HD 207869, HD 77110, HD 61986, HD 111777, HD 199288, HD 22879, and HD 56274. For reference, the posterior distributions for the orbital periods for these stars are shown in Fig. 4.35.

As mentioned above, the posterior ratio criterion is not the only test for a planet detection. And it is clear in Fig. 4.35 that the orbital period posteriors are not constrained for most stars. But note that in Fig. 4.35 we show the joint posterior for the orbital periods of all three Keplerian signals considered within the model. In order to apply criterion 2, we must instead consider the marginal posterior for P_1, if there is evidence for one Keplerian signal, or the joint posterior for P_1 and P_2 if there is evidence for two signals. These are the posteriors that we require to be constrained.

In order to improve our estimates of these posteriors, we ran a new analysis for 10 of these 11 targets. We consider N_p fixed to 1 for all stars except HD 199288, where we fix N_p to 2 instead, since that is the value supported by the posterior for N_p. We leave out HD 56274, which shows a typical posterior distribution for N_p that is caused by stellar activity contaminations (compare it to HD 41248 and HD 175607 in Fig. 4.34) and will be discussed in more detail below.

The new posteriors for the orbital periods are shown in Fig. 4.36, which can be compared to Fig. 4.35. Because these new models have a fixed value of N_p, we can directly apply criterion 2. and assert whether the posteriors for the orbital periods are unimodal and constrained.

It is easy to see that for HD 71685, HD 197197, HD 77110, and HD 111777, the posteriors shows severe multimodality. The causes for this multimodality are not completely clear, but we guess that the presence of long gaps without observations might be causing these strong aliases. HD 71685, for example, has only 32 RV observations and shows a gap in the observations of almost 7 years. For HD 206998, the posterior shows more than one mode, but they are relatively isolated. Nevertheless, its posterior for N_p only provides evidence for one significant Keplerian signal.

There are also some examples of significant signals appearing close to one year or its harmonics (shown as red dashed lines in Fig. 4.35 and Fig. 4.36). Note that this is still the case, even though the RVs of specific stars were re-derived with a custom mask. In principle, the mask avoids all the spectral lines that cross the HARPS CCD stitchings, and should therefore correct any signal caused by this effect. In the case of HD 207869, the one planet solution with period close to 180 days has a very high eccentricity (close to 0.9) and it is quite clearly not being caused by a planet. But for HD 22879, it is not as clear if we should take the RV signal at $\sim$330 days as a bona-fide planet detection.

Finally, there are a few stars for which the posterior distributions show clear and well-defined peaks. HD 126803 and HD 61986 have clear peaks in both analysis, with fixed or varying N_p. Also the results for HD 199288 point to the detection of 2 significant Keplerian signals, and the posterior with fixed $N_p = 2$ presents two isolated modes.

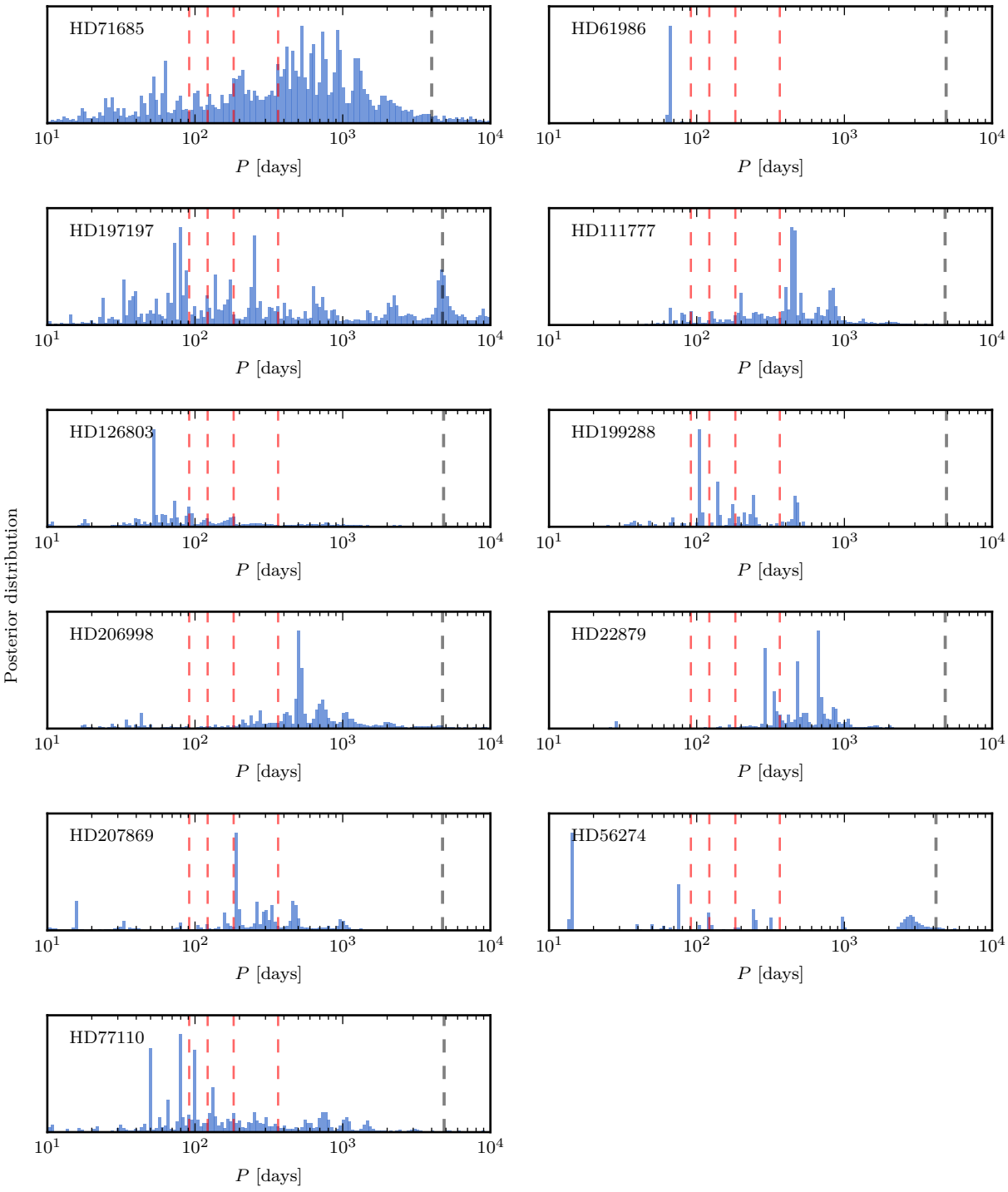


Figure 4.35: Posterior distributions for the orbital periods, for the stars which pass the posterior ratio criterion. The dashed grey line marks the timespan of the RV observations and the dashed red lines are at 365 days and the first harmonics.

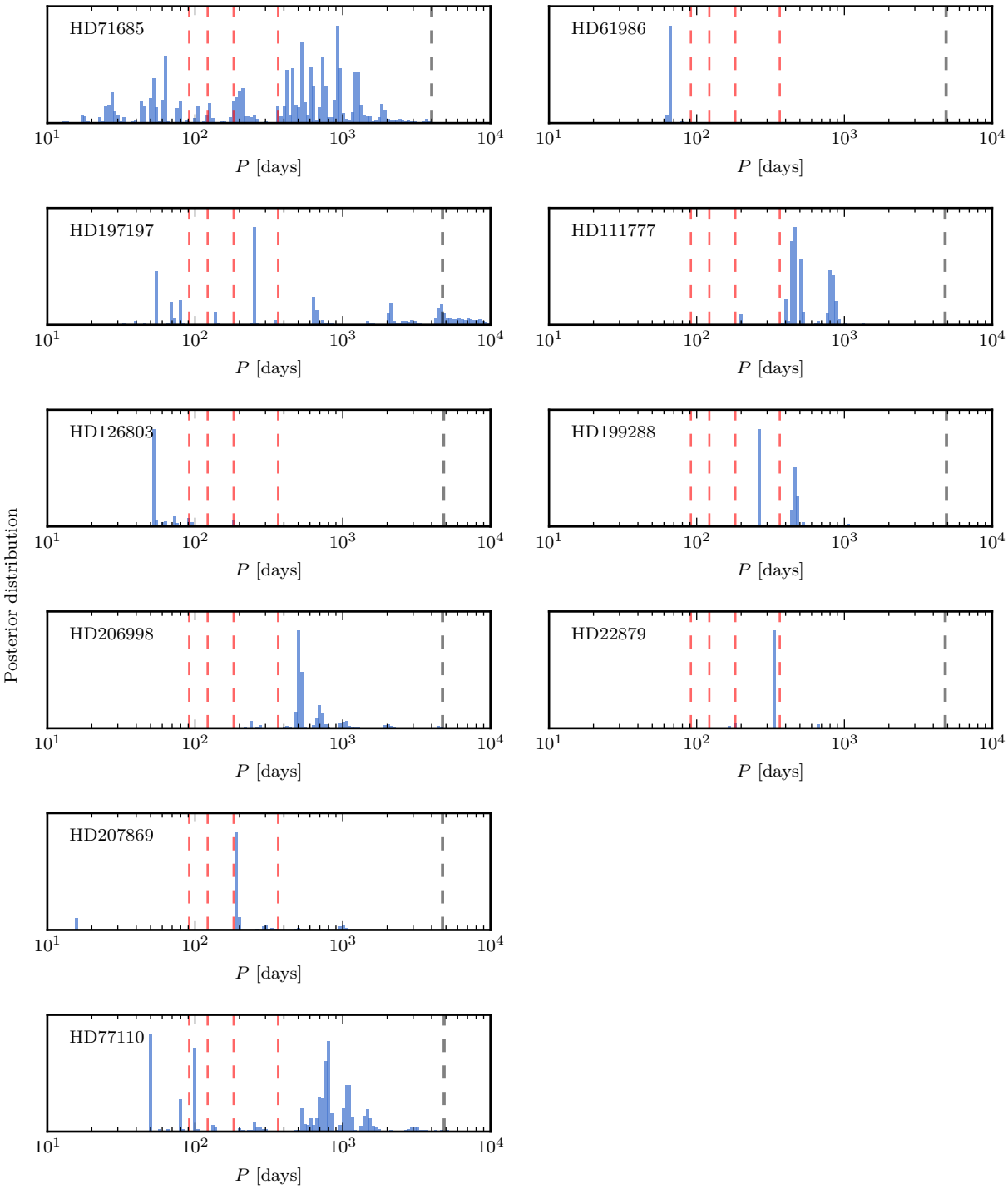


Figure 4.36: Posterior distributions for the orbital periods, from analyses with fixed N_p. The dashed lines are the same as in Fig. 4.35.

Therefore, these three stars pass our criterion 2.

In order to try and summarize these results, the following section discusses what I consider to be the bona-fide planet detections for HD 126803, HD 61986, and HD 199288. Later, we discuss the signals that can only be considered candidates with the available data, for HD 206998 and HD 22879.

4.3.7 *Planet detections*

HD126803 The analysis of this star's 91 RV observations provides decisive evidence for the presence of one Keplerian signal ($\mathcal{B}_{1,0} = 190.1$), with the highest peak in the orbital period posterior being around 50 days. Our estimates for the orbital parameters of this planet are $P = 53.1^{+39.8}_{-0.8}$ days, $K = 1.74^{+0.59}_{-0.44}\,\mathrm{m\cdot s^{-1}}$ and $e = 0.27^{+0.34}_{-0.19}$, calculated using the median, 16% and 84% percentiles of the marginal posterior distributions. These parameters correspond to a planet with 8.3 $M_\oplus$ orbiting at 0.25 AU from its host.

HD61986 This star has been observed 155 times over 89 nights. The complete timespan covers an impressive total of 13 years, even though most observations were obtained recently after the HARPS fibers change. From the posterior distribution (with $N_p = 1$) we obtain the following orbital parameters: $P = 64.74 \pm 0.16$ days, $K = 1.72 \pm 0.27\,\mathrm{m\cdot s^{-1}}$ and $e = 0.18^{+0.21}_{-0.13}$. These values correspond to a planet having a mass $m_p \sin i = 9.46^{+1.53}_{-1.49}\ M_\oplus$ and orbiting at 0.3 AU from the host star. The phase curve of the maximum likelihood solution is shown Fig. 4.37, together with the GLS periodograms of the original data and the residuals after subtraction of the planetary signal. The periodogram of the RV residuals shows one peak at 10 days, but its power is below the 0.1% FAP line.

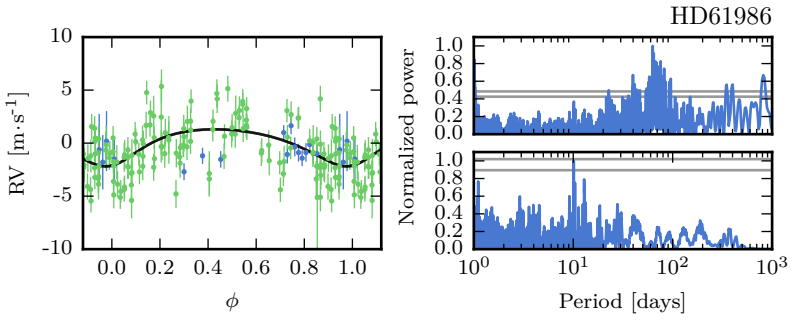


Figure 4.37: Maximum likelihood solution in the 1-planet model, for HD 61986. The left panel shows the phase curve of the planet signal, with RV observations before and after the fibers change shown in blue and green, respectively. The right panels show the GLS periodograms of the original data and the residuals. Horizontal lines mark the 1% and 0.1% FAP levels.

HD199288 Out of the total 195 observations for this star, we removed one clear outlier (at more that 10 sigma from the mean RV) and 10 points that were obtained during the commissioning period for HARPS new fibers. This leaves 184 RV observations, with a timespan of more than 13 years. Our analysis suggests the presence of two Keplerian signals, and the two periods are visible in Fig. 4.36, obtained with N_p fixed to 2. The maximum likelihood solution from this analysis is shown in Fig. 4.38. From the two modes in the posterior, we obtain for the first planet: $P = 260.29^{+0.45}_{-0.52}$ days, $K = 2.34 \pm 0.28\,\mathrm{m\cdot s^{-1}}$,

and $e = 0.81^{+0.45}_{-0.52}$, corresponding to a companion with a mass of $11.9^{+1.57}_{-1.39}$ $M_\oplus$. For the second planet: $P = 458.05^{+18.29}_{-7.74}$ days, $K = 1.77^{+0.27}_{-0.23}$ m·s^{-1}, and a lower but still significant eccentricity of $0.35^{+0.12}_{-0.14}$. These parameters result in a planetary mass of 17.42 ± 2.5 $M_\oplus$. This solution is considered a planet detection, but it will be important to test the dynamical stability of the multiple planet system (see, for example, Goździewski et al., 2006).

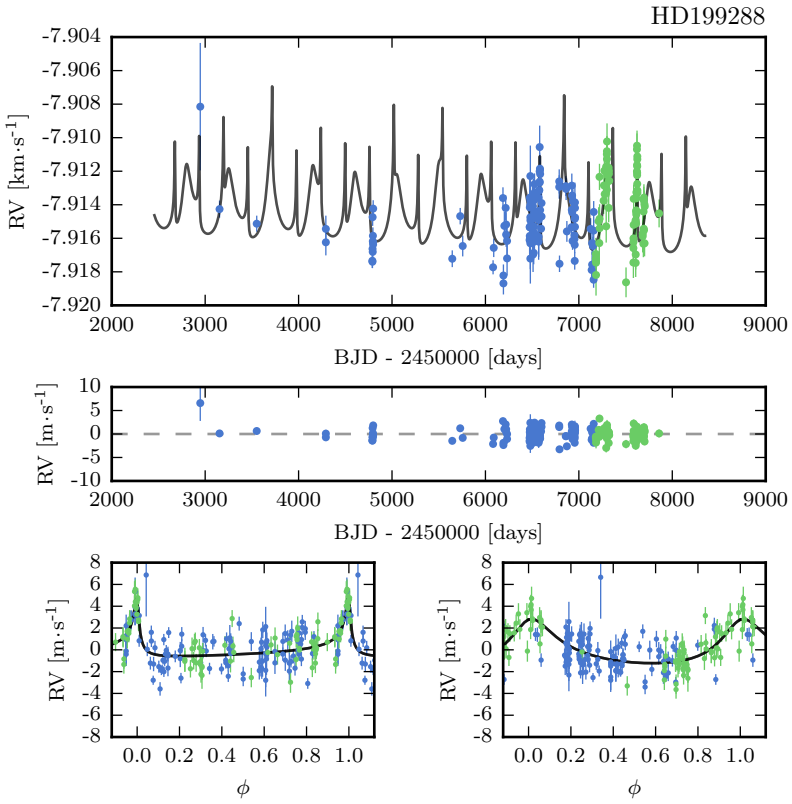


Figure 4.38: Maximum likelihood solution for HD 199288, obtained from the analysis with $N_p = 2$. Observations before and after the fibers change are color-coded. The middle and bottom panels show the residuals and phase curves for the two signals, respectively.

4.3.8 *Planet candidates*

With the RVs of HD 206998 and HD 22879 we could not obtain secure planet detections. Therefore the following solutions are taken as candidates that should be confirmed with additional data.

For the candidate planet around HD 206998, selecting the main peak in the posterior distribution for the orbital period, we obtain $P = 510.68 \pm 12.75$ days and a significant eccentricity ($e = 0.33 \pm 0.21$). With a semi-amplitude around 2.90 m·s^{-1} this corresponds to a planet mass of $30.64^{+8.73}_{-7.16}$ $M_\oplus$.

Our analysis of HD 22879 suggests the presence of one significant Keplerian signal, but its period is very close to one year, making it difficult to claim a planet detection. From the posterior distribution with N_p fixed to 1, we obtain $P = 335.42^{+1.86}_{-1.71}$ and an amplitude of 1.4 ± 0.22 m·s^{-1}. The eccentricity is also high: 0.35 ± 0.12. If indeed caused by a planet, it would have a mass of 12.29 ± 1.94 $M_\oplus$.

4.3.9 HD 56274: *a stellar magnetic cycle*

The ensemble analysis of the RV observations for HD 56274 points to the significant detections of two Keplerian signals (see Fig. 4.34). This star has been observed 235 times on 165 individual nights. The observations cover a timespan of over 11 years. Since the raw RVs of HD 56274 showed signals at one year and its harmonics, they were re-derived with a custom mask, as explained in section 4.3.5. The ensemble analysis presented above used this corrected data.

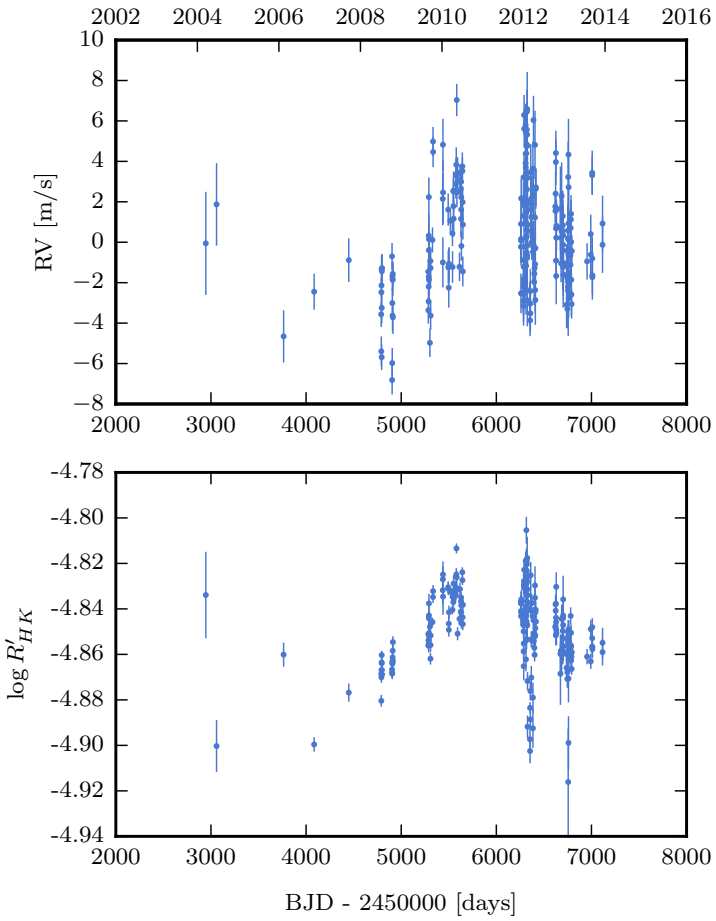


Figure 4.39: Timeseries of the RVs and the $\log R'_{\rm HK}$ measurements for HD 56274.

From the timeseries it is already clear the the RVs show long-period variations, which are in phase with the $\log R'_{\rm HK}$ index (see Fig. 4.39). This suggests the presence of a magnetic cycle, analogous the solar 11-year cycle (see for example Lovis et al., 2011a; Robertson et al., 2013; Johnson et al., 2016). The period associated with this magnetic cycle is already visible in the (last) posterior distribution in Fig. 4.35, which shows a peak around 3000 days. In the model used for the ensemble analysis, this RV variation is being modelled as a Keplerian.

A more careful analysis provides an estimate for the period of the magnetic cycle of $10.48^{+3.80}_{-1.87}$ years.

From the $\log R'_{\rm HK}$ measurements, and using the calibrations of Noyes et al. (1984b) and Mamajek and Hillenbrand (2008), we derive a rotation period for HD 56274 of 13.32 ± 1.02 days and 13.28 ± 2.18 days, respectively (see Table C.1). The main peak in HD 56274's posterior (Fig. 4.35) is actually at this period, which means that the

RVs also show variations caused by activity at shorter timescales.

Because of the impact of stellar activity, a new analysis of the RVs of HD 56274 was carried out with the model of section 3.3, which includes the GP component. The default priors were used (see Tables 2.1 and 3.1), but we set $N_{p,\max}$ to 5. This analysis is thus similar to the one carried out for HD 175607 and HD 41248.

The resulting posterior distribution for N_p is in Fig. 4.40, and shows that, with the analysis that includes the GP component, we no longer detect significant Keplerian signals for HD 56274. The ratio of the posterior probabilities for $N_p = 0$ and $N_p = 1$ is only 5.9, which is well below the 150 required by our detection criterion.

The joint and marginal posteriors for the GP parameters and the extra white noise s, are shown in Fig. 4.41. Both s and η_1, the amplitudes of the extra white noise and of the GP covariance are well constrained by the data. The remaining parameters shows less well-behaved or multimodal posteriors. Nevertheless, note that the clear peak in the posterior for η_3 is close to the estimated rotation period of HD 56274. From this analysis, our estimate for the rotation period, obtained from the RVs, is $17.80^{+15.84}_{-3.61}$ days.

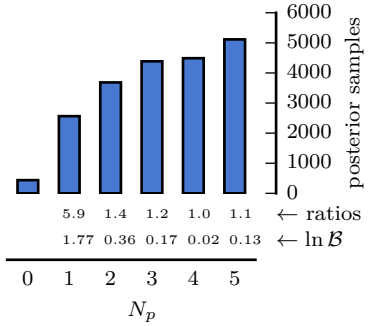


Figure 4.40: Posterior distribution for the number of planets, for HD 56274.

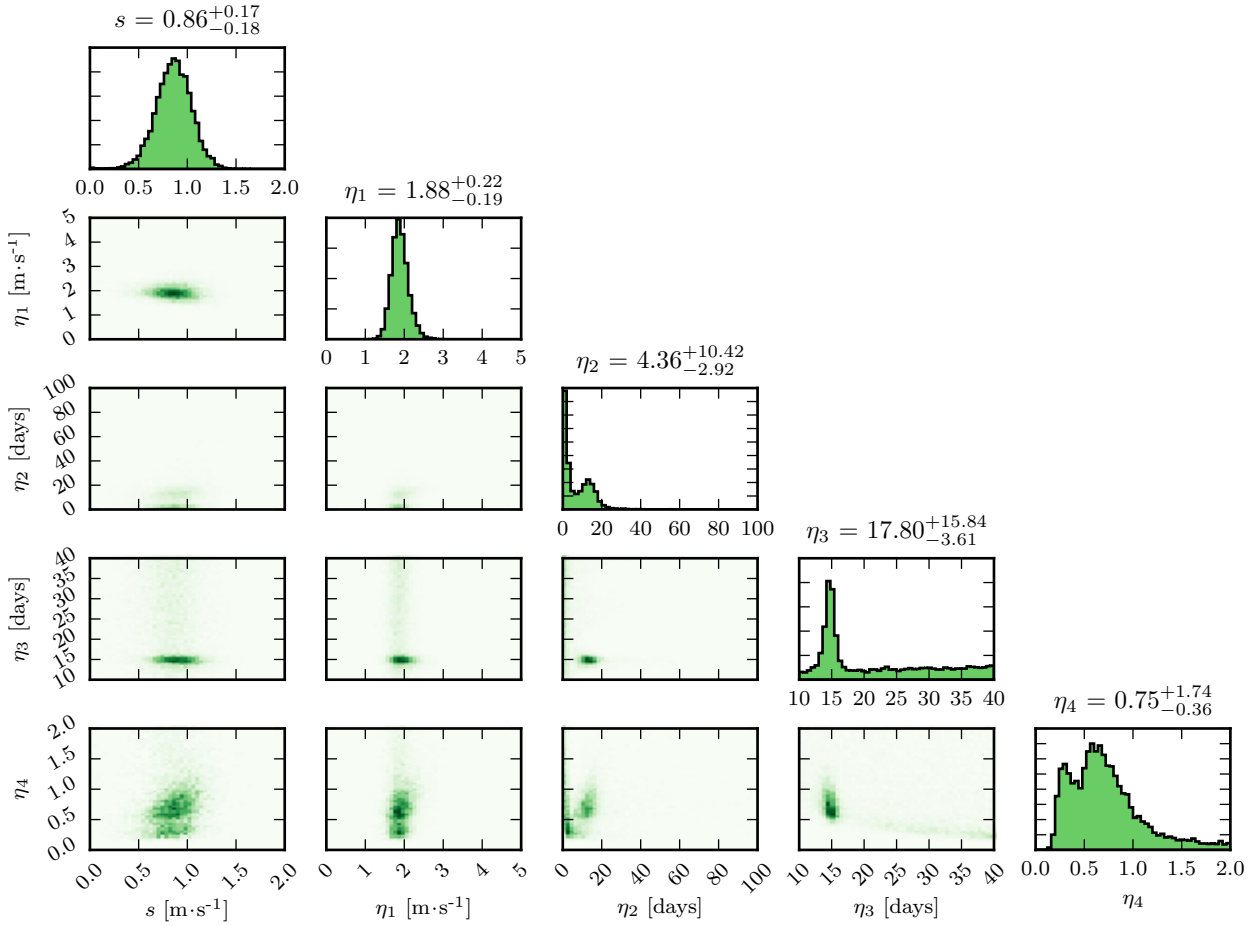


Figure 4.41: Marginalized and joint posterior distributions for the GP parameters, from the analysis of HD 56274. The value at the top of each distribution represents the median and 16% and 84% quantiles of the posterior samples.

It seems that our RV model that incorporates the GP component is able to model the short-period variations that are linked to stellar

activity. But note that the model was not designed to model magnetic cycles, such as the one we detect for HD 56274. In practice, this means that the quasi-periodic kernel is probably not the most adequate, even if it still manages to describe some part of the RV variations.

Usually, these variations caused by magnetic cycles are modelled using sinusoidal or Keplerian signals (e.g. Lovis et al., 2011a; Johnson et al., 2016), or considering linear correlations between RVs and $\log R'_{\rm HK}$ (e.g. Dumusque et al., 2012). Therefore, we can think of possible extensions of our model to consider such long-term variations. We could include a long-period sinusoidal signal to be interpreted as the magnetic cycle; we could consider a correlation term with the $\log R'_{\rm HK}$, as in Tuomi et al. (2014a), for example; or we could add an (periodic) kernel to the covariance matrix of the GP. These ideas could certainly improve our analysis of HD 56274 in particular.

4.4 OCCURRENCE RATES

Having run our planet detection pipeline on all the stars in the sample, the aim of this Section is to calculate the frequency of planets orbiting metal-poor stars. Obviously, any constraint we can place on this frequency is based on the range of metallicity of the stars in the sample, and on which planets we can detect with our RV observations. In section 4.4.1, the detection limits will first be calculated for each star and then combined to yield the completeness of our survey. Afterwards, in section 4.4.2, I will derive the planet frequency and place the final constraints on the occurrence of low-mass planets.

4.4.1 *Detection Limits*

For every given star in our sample, the limited number of RV observations and their specific sampling, can be translated into a sensitivity curve for the detection of planets with a given mass and period. In other words, using the actual observed RVs, we want to calculate an upper limit on the signal amplitude (or the planetary mass) that an undetected companion can have, as a function of orbital period.

This calculation of RV *detection limits* is commonly done in the analysis of RV surveys. Two main approaches have been considered: either based on χ^2 and F-tests as in Lagrange et al. (2009) and Sozzetti et al. (2009), or based on a periodogram analysis, as in Cumming et al. (1999), Endl et al. (2001), Cumming (2004), Mayor et al. (2011), and Mortier et al. (2012), for example. Usually, which of these two approaches is used depends on the average number of RV measurements and on whether the typical sampling is taken as being prone to a periodogram analysis.

Here I will follow closely the method described in Mortier et al. (2012), and already implemented in Faria et al. (2016b) for the analysis of a sub-sample of stars from the LP. This method is based on the injection of planetary signals in the observed RVs, followed by a periodogram analysis.

The procedure is as follows. We explore all orbital periods in the range from 1 day to 5000 days, and semi-amplitudes up to $10\,m{\cdot}s^{-1}$. To speedup the computations, instead of considering a grid of semi-amplitudes, a binary search algorithm is used to bracket the undetected amplitude more efficiently. For each orbital period, each semi-amplitude and ten linearly-spaced phases, we inject a sinusoidal signal with those parameters in the observed RVs. Then we compare the periodogram power of the injected signal with the 1% FAP level of that period in the original dataset. If, for all phases, the former is higher, we consider the planet to be detected. We convert the semi-amplitude to planetary mass using the stellar masses derived for each star (see Table C.1).

In other words, we assert that an injected signal is detected (and thus *detectable* in this dataset) if the periodogram power at that period is higher than the 1% FAP level of the original dataset's periodogram.

It is important to note that this method assumes that the original dataset (in which the mock planets are injected) only contains uncorrelated noise. For most stars, this assumption is valid as we have not detected any significant planetary signal. For the giant planet hosts (4.3.2) and HD 126803, HD 61986, and HD 199288, we subtracted the adopted orbital solution from the data before calculating the detection limits. For HD 41248, HD 175607, and HD 56274 we subtract from the data the model predictions calculated with the mean posterior values, which includes the GP activity model.

For other stars where we do not constrain the orbital parameters of the candidate planets, we analyse the original RV timeseries making the asusmption that they contain only noise. Therefore, for these stars, the detection limits estimated here can be taken as conservative.

It is also very important to note that, since we use sinusoids as our model for the data, the detection limits we derive are strictly valid only for circular orbits. However, this has been shown not to be a severe limitation, because the periodogram still gives a good estimate of the orbital period even for eccentric orbits (see,e.g., Cumming et al., 1999, and references therein). Another known caviat of this analysis is that only single planet systems are simulated, therefore assuming that the detection of one planet does not affect that of another.

Fig. 4.42 shows a few examples of detection limits calculated with this procedure. A few stars were chosen, with representative curves. HD 128340 and HD 31128 are the stars with the least and the most RV observations, respectively: it is clear that only 5 observations are not enough to constrain the presence of low-mass planets, even at very short periods; but with 264 points, the $1\,m{\cdot}s^{-1}$ limit is almost reached in a large range of periods, even for the most metal-poor star in our sample. HD 196892 and HD 22879 are shown to represent a typical curve and the lowest detection limits obtained in our sample.

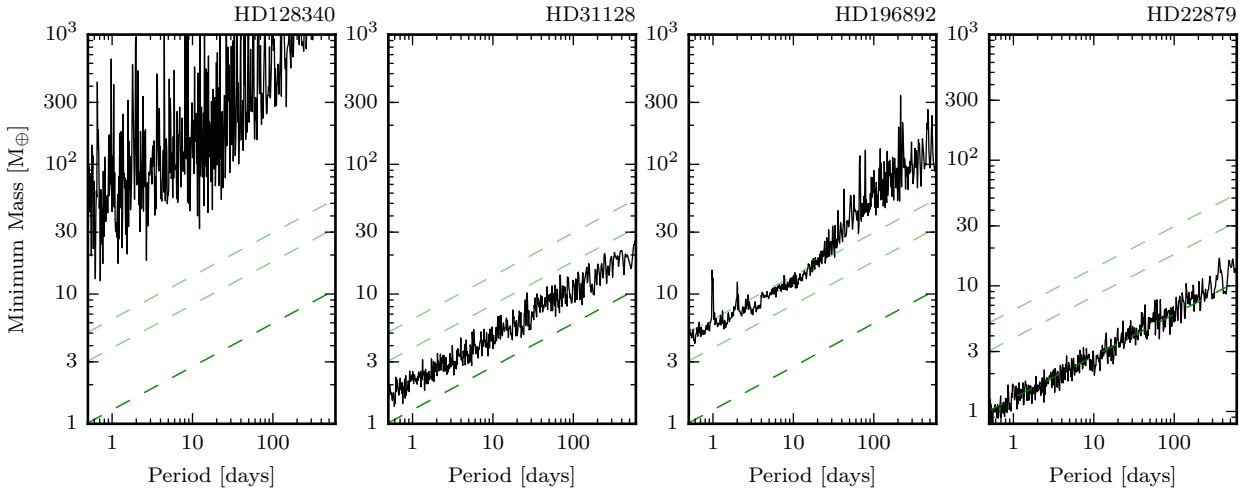


Figure 4.42: Representative detection limits from four stars in the LP. The dashed green lines show a circular planetary signal with an RV semi-amplitude of 1, 3 and 5 $m \cdot s^{-1}$ (from bottom to top).

4.4.2 *Frequency of planets*

To calculate constraints on the frequency of low-mass planets from our survey, we need to obtain the overall limits of the survey by combining the detection limits for individual stars. We will calculate the *completeness* function, $C(m_p, P)$, as the fraction of stars in our sample with sufficient measurements to detect or exclude a planet, at a given period and minimum mass (see e.g. Mayor et al., 2011, which, for simplicity, is hereafter referred to as M11). This function is obtained by directly combining the individual detection limits, and is represented in Fig. 4.43.

The completeness function shows the same behaviour as the individual detection limits. It is interesting to see that it shows lower values at periods of 1, 2 and 365 days. These arise naturally from the sampling of the individual timeseries and from problems such as that described in section 4.3.5 (see also Dawson and Fabrycky, 2010). Otherwise and as expected, $C(m_p, P)$ is a decreasing function of orbital period for fixed planet mass, and an increasing function of planet mass for fixed orbital period.

In Fig. 4.43 we have also included the planet detections from the previous sections. The two planets in systems (HD 171028 and HD 199288) are connected. For reference, we also show the masses of some Solar System planets, arbitrarly placed at 1500 days. The mass – period regions where we estimate the planet frequency are marked as dashed lines (see the discussion below).

Using the completeness function and the number of planets found in the sample, we can analyse statistically the planet frequency. This calculation is usually performed in one of two ways, either based on the binomial distribution or by considering a parametric fit to the fraction of planets detected. The very low number of planet detections forces us to consider only the first method.

The method is well described in Burgasser et al. (2003) and Mortier et al. (2012) and we used it also in Faria et al. (2016b). The binomial distribution provides the probability of having k detections in a

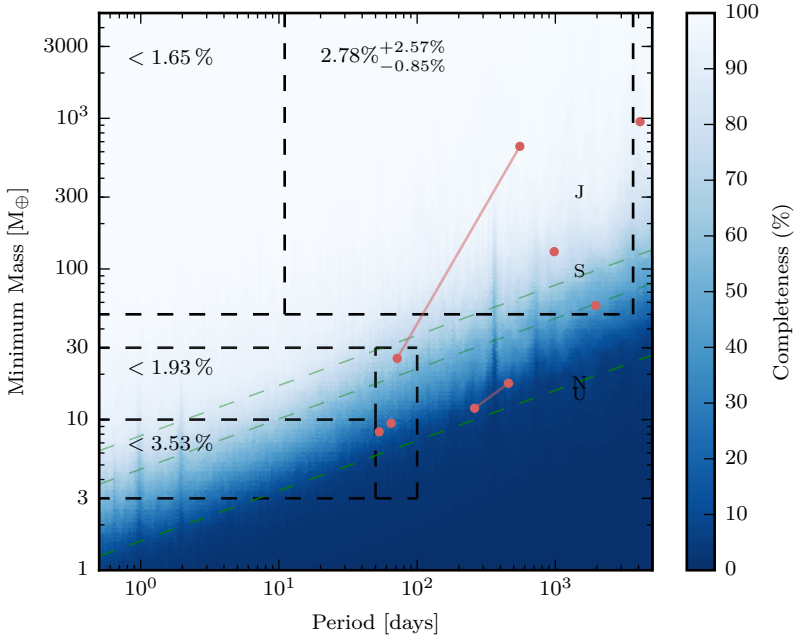


Figure 4.43: Completeness of the metal-poor survey. The colormap shows $C(m_p, P)$ as a function of orbital period and minimum mass. Planet detections are identified as red circles, connected by a line when they correspond to multiple systems. Dashed black lines delimit the regions where we calculate the planet frequency and our estimates for f_p are also shown. The masses of some Solar System planets are marked at an arbitrary period of 1500 days.

sample of size N, given a true planet frequency f_p:

$$p(k \mid N, f_p) = \frac{N!}{k!(N-k)!} f_p^k \, (1 - f_p)^{N-k}. \tag{4.5}$$

This equation can represent a likelihood function, when seen as a function of f_p for given values of k and N. Applying Bayes' theorem with a uniform prior for f_p between 0 and 1, we obtain a posterior distribution for f_p:

$$p(f_p \mid N, k) = \frac{p(k \mid N, f_p)\, p(f_p)}{\int_0^1 p(k \mid N, f_p)\, p(f_p)\, df_p} \tag{4.6}$$

$$= \frac{f_p^k \, (1 - f_p)^{N-k}}{\frac{k!(N-k)!}{(N+1)!}} \tag{4.7}$$

$$= (N + 1)\, p(k \mid N, f_p). \tag{4.8}$$

Therefore, the number (k) of planets detected in a given sample (of size N) places a constraint on the probable values of the planet frequency, which can be expressed by the mode and the 68% credible interval of the posterior distribution. In practice we estimate this credible interval by numerical integration of the posterior.

In Faria et al. (2016b), we studied a subsample of stars from the LP which had more than 75 measurements at that time. Those 15 stars had already enough observations so that the detection limits were below the threshold of $(m_p \sin i, P) = (10\, M_\oplus, 50\, \text{days})$, therefore allowing for a comparison with the results from M11. In that case, we could use directly $N = 15$ in Eq. (4.8).

Now with the complete sample, only a fraction of stars reach that threshold. The "effective" sample size N has to be calculated based

on the completeness function $C(m_p, P)$. We will use the average completeness in a given region of the mass – period space, and multiply that value by 109 stars to obtain N. Plugging that value in Eq. (4.8) will provide a constrain on the frequency of planets up to a given period and down to a given (minimum) mass.

Table 4.5 lists our results, in comparsion with the planet frequencies obtained by M11. We consider the same regions in parameter space that were used by M11 and also considered in Lovis et al. (2008). In the regions where we have planet detections, our estimates for the planet frequency correspond to the mode and the 68% credible interval of the posterior distribution for f_p. If there are no planets detected, we list only the 68% upper limit. Some of these values are also shown in Fig. 4.43, next to their corresponding region in the parameter space. In this way, Fig. 4.43 can be more directly compared with the Figures 6, 7 and 9 from M11.

Mass	Period	Planet frequency solar-metallicity	Planet frequency metal-poor	Comments
$> 50\ M_\oplus$	< 10 years	13.9 ± 1.7 %	$2.78\%^{+2.57\%}_{-0.85\%}$	Gas giants
$> 50\ M_\oplus$	< 11 days	$0.89 \pm 0.36\%$	$< 1.65\%$	Hot gas giants
$< 30\ M_\oplus$	< 100 days	47.9 ± 8.5 %	$6.38\%^{+5.55\%}_{-1.98\%}$	Low-mass
$< 30\ M_\oplus$	< 50 days	38.8 ± 7.1 %	$< 3.46\%$	Low-mass, short periods
$10 - 30\ M_\oplus$	< 50 days	$11.1 \pm 2.4\%$	$< 1.93\%$	
$3 - 10\ M_\oplus$	< 50 days	$16.6 \pm 4.4\%$	$< 3.53\%$	

Table 4.5: Frequency of stars with planets in various regions of the $m_p \sin i$ – period plane. The values for the solar-metallicity stars are from M11, based on the HARPS and CORALIE samples. Our estimates are obtained with the complete metal-poor sample.

In general, our estimates for the frequency of planets are significantly lower than the results on solar-metallicity stars. Starting with the giant planets, we are not able to constrain efficiently the frequency of hot gas giants, due to the smaller size of our metal-poor sample (Mayor et al.'s combined sample had 822 stars from the HARPS and CORALIE surveys). For these planets, with masses larger than 50 $M_\oplus$ and periods shorter than 11 days, our results are compatible with those of M11. For the longer period gas giants, with orbital periods up to 10 years, we find a significantly lower frequency of $2.78\%^{+2.57\%}_{-0.85\%}$, compared to almost 14% for the solar-metallicity stars. Note that the 10-year period limit means that only HD 11397 b, HD 171028 b and HD 181720 b are included in the calculation, leaving out the giant planet companion to HD 190984. Including this planet would change the planet frequency to $3.70\%^{+2.76\%}_{-1.08\%}$, still significantly lower than the solar-metallicity value.

For the lower mass planets we also obtain substantially lower frequencies, in all mass and period regions. For planets with up to 30 $M_\oplus$ and periods shorter than 100 days, M11 reports a planet frequency of almost 50%, while we obtain an estimate lower than 12% at the one-sigma level. In the shorter period regions, our results are even more discrepant. This suggests a very significant role of metallicity in shaping the low-mass planet population. We will discuss these results in the final Chapter 5, below.

5

Discussion and Conclusions

At the onset of this Thesis, I was faced with an open question: *how common are low-mass exoplanets around metal-poor stars?* The three chapters above present my efforts towards answering this question. In the present Chapter, I reflect on the conclusions from my work, and the new questions that can bear from it.

As part of an HARPS RV survey, we have been following a sample of metal-poor stars for more than a decade. With this program, we have gathered a tremendous ammount of data, both in actual quantity (7298 individual spectra) and in scientific relevance. HARPS is one of the most precise spectrographs in the world, and this is (to our knowledge) the only survey focused on metal-poor stars, capable of detecting planets with masses close to that of the Earth.

* ◯ *

To fully exploit all these data, and building upon the pioneering work of Brewer (2014) and Brewer and Donovan (2015), I have presented a general method for the detection of exoplanets in RV data. This method uses the power of Bayesian inference to provide estimates for the planets' orbital parameters, *together* with the number of confidently detected planets in a given dataset.

In a recent account of the state-of-the-art in the RV search for exoplanets, Fischer et al. (2016) mentioned a "dramatic shift over the past two decades", as Bayesian methods started to be applied almost ubiquitously in the field. Ford (2005) and Gregory (2005) led the initial revolution, applying MCMC algorithms to estimate the posterior distributions of the parameters of interest, such as the planets' orbital periods and masses. Over the years, the need for rigorous model comparison to quantify the evidence for planet detections became apparent (e.g. Feroz et al., 2011b; Tuomi and Jones, 2012), and sophisticated methods were developed towards this goal.

The model presented in Chapter 2, and its extension, described in Chapter 3, is state-of-the-art in these efforts. It has three main ingredients: an efficient sampling algorithm, capable of estimating the evidence of a model; a framework that allows for the number of planets N_p to be included as a free parameter; and a set of general-purpose priors which can be applied to a number of different systems.

The sampling algorithm, based on the Diffusive Nested Sampling algorithm Brewer et al., 2011, see also Appendix C), is capable of sampling from very complicated posteriors, and to do it efficiently. It has

virtually no tuning parameters, does not depend on starting values, and can sample from severely multimodal distributions, such as those often found for the orbital period in the analysis of an RV dataset. It further allows for the estimation of the Bayesian evidence integral (see Eq. 2.76), which is the central quantity in model comparison.

The birth-death (or reversible-jump) framework which allows for N_p to be a free parameter (Brewer, 2014), means that the comparison between different models is done in the most straightforward way possible. This method does not require multiple runs with varying N_p (Feroz et al., 2011b), shows good scaling with the value of $N_{p,\max}$ and with the number of observations, and provides estimates for all the parameters of each of the N_p signals (which may be planets).

Our threshold for identifying a significant detection, based on the Jeffreys (1961) scale of evidence, is an important point in all the analyses. It can be seen as a conservative criterion on itself (see Tuomi et al., 2014b), especially when including also the two criteria from section 4.3.6. The reader may be left wondering (as the author definitely is) how many of the known exoplanets discovered in RV datasets would survive the scrutiny of the posterior ratio criterion. Answering this problem is beyond the goals of this Thesis, but is an avenue to be explored in the future.

As far as the set of priors, they are nothing less than remarkable. It is worth noting that virtually every analysis I presented in this Thesis was carried out with the same priors.[1] This attests both to the efficiency and general-purpose of the algorithm and, since the priors are mostly identical to those considered by Brewer and Donovan (2015), to the importance of their work.

On the other hand, Bayesian model comparison is (by construction) sensitive to the priors one considers. This sensitivity is an asset, not a shortcoming, but it can mean that some of our results would be different had we considered a different set of priors. In particular, the effect of the hyperparameters $\boldsymbol{\alpha}$ on the detection criterion is not studied completely. One could compare the volume of the parameter space when $N_p = 0$ to that when $N_p = 1$, for example. The latter is much larger when including $\boldsymbol{\alpha}$. We expect that this leads to a more conservative model than if $\boldsymbol{\alpha}$ were not included.

[1] We did change the priors in a few cases, in order to illustrate some points, but the results for simulated data, the ELODIE data of 14 Her, the HARPS data of HD 10180 and CoRoT-7, as well as all the data from the metal-poor program, were all obtained with the same priors.

Remember that $\boldsymbol{\alpha} = \{\mu_P, \omega_P, \mu_K\}$, corresponding in multiple planet systems to a mean orbital period, a spread in orbital periods, and a mean semi-amplitude, respectively.

* ◯ *

The RV community is currently facing the challenge of separating real planetary signals from intrinsic velocity fluctuations of the stars (Fischer et al., 2016; Dumusque et al., 2017). The manifestations of stellar oscillations, granulation and magnetic activity are observed as complex RV signals, sometimes with very similar amplitudes and periodicities to those of planets (see Chapter 3). As we search for lower mass planets, this limitation becomes more and more apparent, making it difficult to detect and correctly characterize these planets (e.g. López-Morales et al., 2016).

To help in solving these problems, I have extended the original

planet detection method, which only considered Keplerian signals, with a model for stellar activity. The model uses Gaussian processes to describe the quasi-periodic RV variations induced by the presence of active regions in the stellar surface (section 3.3.1). This is a statistical model, but it is also physically-motivated. The quasi-periodic kernel that we use (see also Haywood et al. 2014; Grunblatt et al. 2015) has four parameters in total, two of which can be related to the properties of the active regions. The stellar rotation period is described by η_3, and the timescale for evolution of the active regions by η_2.

The application of this model to CoRoT-7 , described in Section 3.4 and in Faria et al. (2016a), revealed a number of interesting results. First, we were able to recover the orbital parameters of CoRoT-7 b and CoRoT-7 c, with results fully compatible to those of Haywood et al. (2014), who had used photometric information from CoRoT. At the same time, we did not find enough evidence for a possible third planet, as proposed by Hatzes et al. (2010), for example. By calculating the posterior distribution for η_3, we also obtained an estimate for the rotation period of CoRoT-7, solely from RV observations.

This extended model has also been applied to some of the stars in the metal-poor sample. In particular, we tested the detection of HD 175607 b, announced by Mortier et al. (2016), and the two possible companions to HD 41248, announced by Jenkins et al. (2013) and Jenkins and Tuomi (2014). In both cases, we do not find evidence for the reported planet detections. For HD 175607, it seems that the restrictive prior used by Mortier et al. (2016) for the orbital period of the planet had an impact on the results of the model comparison. For HD 41248, on the other hand, it is not completely clear what difference between our model and that of Jenkins and Tuomi (2014) could explain the different results (but see the discussion in Santos et al. 2014). This shows that detecting exoplanets in the RVs of active stars is definitely not yet a solved problem.

* ◯ *

The results from Chapter 4 are the most important scientific contribution from my Thesis, in tandem with the more technical or methodological previous chapters. After analysing the sample of metal-poor stars that are part of our HARPS RV survey, I estimated the frequency of planets with different properties in this low-metallicity regime and can now place constraints on the models of planet formation.

From the sample of 109 stars, with derived spectroscopic metallicities in the range $-1.3 < [\mathrm{Fe/H}] < -0.05$, we were able to find:

- six long-period giant planets (three of which were already known) HD 11397 b and c, HD 197083 b, HD 171028 b, HD 181720 b, and HD 190984 b (see sections 4.3.1 and 4.3.2).
- three brown-dwarf companions HD 107094 B, HD 144589 B, and likely HD 93351 B.
- five lower-mass planets for HD 126803, HD 61986, HD 199288 b and c, and HD 171028 c.

Regarding the giant planets, we compared our estimated planet frequency to the value presented by Mayor et al. (2011), and found the former to be significantly lower (see Table 4.5). But note that, because of the giant-planet – metallicity correlation, the giant-planet frequency is a strong function of metallicity. The value of 13.9% reported by Mayor et al. (2011) is the average frequency calculated for their complete sample. Sousa et al. (2011b) report a giant-planet frequency of 3.77% for metallicities between -0.6 and -0.5 and an increase to 25.58% for higher metallicities, within $[+0.3, +0.4]$. Our results are consistent with a constant giant-planet frequency for stars with sub-solar metallicities, as found by Mortier et al. (2012).

With the remaining detections, we found that low-mass planets in short periods (< 100 days) are uncommon in the metal-poor regime, when compared with the solar-metallicity regime. Our estimates for the frequency of stars with planets (see section 4.4.2) are significantly and many times smaller than those found by Mayor et al. (2011).

A number of factors can contribute to this discrepancy, some of which I already discussed above. First, our criterion for planet detection is conservative and a completely self-consistent analysis would require that we analyse the stars from the HARPS and CORALIE solar-metallicity samples with our method. This would allow for a more fair comparison of the planet frequencies. However, note that a very large fraction of the solar-metallicity planets would have to be "disproved" in order to make the frequencies compatible.

Another possibility, perhaps more interesting, is that metal-poor stars could host a larger number of low-mass planets, or show stronger stellar activity contaminations, than more metal-rich stars. If this is the case, detecting these planets with a limited amount of RV observations would be more difficult. Our results are somewhat sensitive to this scenario, because we calculated detection limits assuming the presence of only one planet and no stellar activity contaminations.

These issues aside, we must still address the implications to planet formation of a lower planet frequency in metal-poor stars. The core-accretion scenario predicts an increased frequency of *giant* planets in metal-rich disks, which gives rise to the observed giant-planet – metallicity correlation (e.g. Ida and Lin, 2004; Santos et al., 2004b). On the other hand, low-mass planets can be formed by core-accretion in disks with a wide range of metallicities (Buchhave et al., 2012; Mordasini et al., 2012; Hasegawa and Pudritz, 2014). But a critical point is that these low-mass planets might orbit at larger orbital periods (e.g. Adibekyan et al., 2013; Courcol et al., 2016).

Low-metallicity disks are known to disperse faster than high-metallicity ones through photo-evaporation (Ercolano and Clarke, 2010). The lifetime of a disk is likely to follow a $\sim 10^{[\mathrm{Fe/H}]}$ dependence (Yasui et al., 2010; Hasegawa and Hirashita, 2014). This means that the disk lifetime at $[\mathrm{Fe/H}] = -1$ is already 10 times smaller than at solar metallicity. One implication is that the low-mass cores formed in metal-poor disks would stop migrating earlier than those planets that are migrating through a more metal-rich disc. As a result, low-

mass planets around metal-poor stars would currently orbit outside of the period range that is probed by RV surveys.

The few planet detections in our metal-poor sample add strength to this emerging picture. Having metallicities of -0.63, -0.61, -0.48, and -0.34 for HD 199288, HD 126803, HD 171028, and HD 61986, respectively, these planet hosts take their place on the lower left corner of Fig. 1.7. The longer orbital periods of these planets (larger than 50 days) fit well within the trend found by Courcol et al. (2016). Our results therefore point to the conclusion that lower-mass planets around metal-poor stars are found at longer orbital periods.

5.1 FUTURE PROSPECTS

As we approach the end of this Thesis, we realize just how much remains to be done! We saw in Chapter 4 that the detection of planetary signals in RV data is still far from being a solved problem. Even with all the prior knowledge that we had to use when setting up our model (which can and should be discussed), and however efficient our methodology proves to be, we cannot detect every planet in every dataset, at the push of a button.

The method described in this Thesis has vulnerabilities, like any other method. Issues related to the sampling of the RV observations, and the signals induced by the instrument itself, are amongst the known vulnerabilities. Many more are currently unknown. And it is remarkably hard to solve these issues with prior information and statistics, so perhaps we should aim to improve the other component of the model: the data themselves.

On one hand, we need to correct for instrumental systematics at the highest possible level of precision (e.g. Pepe and Lovis, 2008). We saw that the correction to the effect of the HARPS CCD stitchings (Dumusque et al., 2015b) was often suboptimal, leaving behind periodic signals, which we think are still related to the motion of the Earth around the Sun. Moreover, in the context of a RV survey, the scheduling of the observations should be adapted as the survey develops, and depending on what signals are detected. Previous research on *adaptive scheduling* (Loredo, 2004; Baluev, 2008; Ford, 2008) has not seen widespread use in the RV community, but it is one clear avenue to mitigate the effects of the time sampling.

The number of new spectrographs coming online in the past months or in the near future is incredible (see Fischer et al., 2016, and references therein). Most of them will deliver RV measurements with precisions of a few $m \cdot s^{-1}$ or even some $cm \cdot s^{-1}$ (Pepe et al., 2014b). Extracting the most information out of these future data is bound to reveal new and exciting properties of exoplanets.

It is also clear that the effects of stellar activity in RV observations are not completely corrected for by the inclusion of GPs in our RV model. We must use our detailed knowledge of stellar activity to inform the statistical models we use. A lot is already known about solar

activity (e.g. Bumba and Kleczek, 1976; Noyes, 1986; Hathaway, 2015) and how it evolves with time. In particular, recent and future projects to measure the Sun-as-a-star RV variations – the solar telescope at HARPS-N (Dumusque et al., 2015a; Phillips et al., 2016), the HELIOS project at HARPS, and other projects (Reiners et al., 2016) – will shed light on the effects of activity when everything else (the planets) is known. Much of this knowledge can hopefully be generalized to other stars, as it has been the case before (Schrijver, 2002).

This is to say that the quasi-periodic kernel that we used to model stellar-induced RV variations is *too general*, because it only takes into account a small part of this knowledge. We only assumed that the RV variations are smooth and quasi-periodic. But the RV variations in a spot-dominated star should look substantially different from those of a plage-dominated star (e.g. Hall et al., 2009; Dumusque et al., 2014a). The ones in a G dwarf should look different from those of an M-dwarf (Bonomo and Lanza, 2012; Robertson et al., 2013). This information needs to be included in a successful model of stellar activity.

Also regarding the presence of low-mass planets orbiting metal-poor stars, we have only scratched the surface. A full understanding of planet formation requires the study of the metal-poor regime, which can provide important tests for the physical processes included in the models (e.g Hasegawa and Hirashita, 2014).

In the future, with exquisite instruments such as the ELT (e.g. Quanz et al., 2015), PLATO (Rauer et al., 2014) or ESPRESSO (Pepe et al., 2014b), we are paving our way towards the detection of other Earths. The methods and know-how developed in this Thesis can and will help in the study of the properties of such exoplanets.

Appendix A: *Diffusive Nested Sampling*

The ability to describe complicated posterior distributions is one of the most important aspects in modern Bayesian analyses. In most interesting cases, an analytical form for the posterior is not available and it must be approximated by numerical methods.

The basic problem is the following: we have a posterior probability distribution $\mathcal{P}(\Theta)$ given by

$$\mathcal{P}(\Theta) = \frac{1}{Z}\pi(\Theta)\mathcal{L}(\Theta) \tag{A.1}$$

where Θ is a parameter vector (which can be of high dimension), $\pi(\Theta)$ and $\mathcal{L}(\Theta)$ are the prior distribution and likelihood, and Z is normalizing constant, the evidence:

$$Z = \int \pi(\Theta)\,\mathcal{L}(\Theta)\,d\Theta. \tag{A.2}$$

We are usually in the situation where Z is not known and so we can only evaluate (more or less easily) a function that is proportional to the posterior distribution. We would like to produce samples from this distribution and, ideally, also to evaluate Z.

The most common algorithm used to generate samples from posterior distributions is a Markov chain Monte Carlo (MCMC), usually paired with a Metropolis-Hastings proposal (see e.g. Geyer, 1992; Berg, 2004; Gamerman and Lopes, 2006, though the literature is incredibly vast). Starting from an arbitrary intial position Θ_i, the next sample Θ_{i+1} is proposed and an acceptance ratio is calculated as $\alpha = \mathcal{P}(\Theta_{i+1})/\mathcal{P}(\Theta_i)$. If $\alpha \geq 1$, the proposed sample is accepted and a new proposal made, otherwise the new sample is accepted only with probability α. If it is rejected, the algorithm proposes a new sample, and so on. A sequence of these steps creates a chain of samples that follow the distribution $\mathcal{P}$.

MCMC can be an incredibly efficient method to sample complicated posterior distributions. However, it leaves the evidence as a secondary quantity, to be calculated by other methods, such as thermodynamic integration (e.g. Lartillot and Philippe, 2006; Calderhead and Girolami, 2009; Gregory, 2011), or as a second step after the MCMC run (e.g. Chib and Jeliazkov, 2001; Perrakis et al., 2014). The Nested Sampling algorithm has been introduced by Skilling (2004, 2006) as a method aimed directly at the calculation of the evidence, and which provides posterior samples as a by-product.

Describing Nested Sampling is relatively easy but understanding *why* it works is not. Let us first detail the steps involved in the method and later attempt at gaining some insight on what it is doing.

Nested Sampling

We start with a population of N particles, $\{\Theta_i\}$, drawn from the prior distribution $\pi(\Theta)$. The value of the likelihood is calculated for each particle. The worst particle – the one with the lowest value of the likelihood – is recorded and removed from the population, being replaced with a new sample drawn again from the prior, but now subject to the constraint that its likelihood must be higher than that of the point it is replacing. It is easy to see that, as this process is repeated, the population moves towards higher values of the likelihood.

Each time the worst particle is recorded, a value $X \in [0,1]$ is assigned to it, representing the amount of prior mass that is estimated to lie at higher likelihood values. After j steps, the prior mass is expected to shrink to

$$\log X_j \sim -\frac{j \pm \sqrt{j}}{N}$$

so that, for a crude implementation, we can set $X_j = \exp(-j/N)$. Assigning these X-values to the particles creates a mapping from the parameter space to the space $[0,1]$, where the prior becomes a uniform distribution and the likelihood is a decreasing function of X (see Fig. A.1). The evidence can then be computed by simple numerical integration of $\mathcal{L}(X)$.

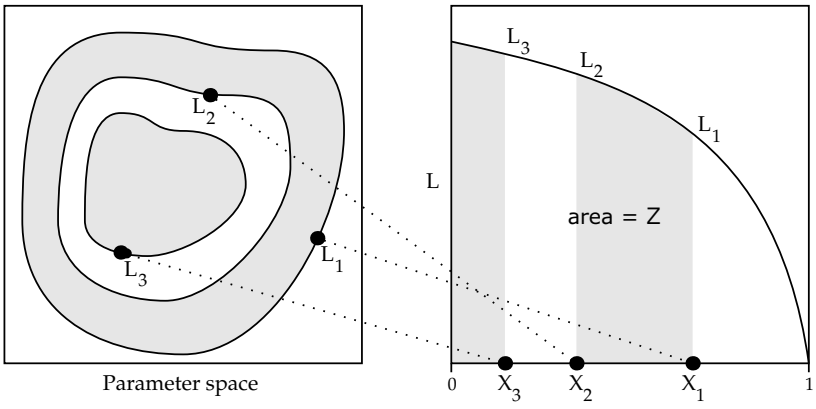


Figure A.1: Visual depiction of the idea behind Nested Sampling: the transformation from a complicated multi-dimensional parameter space to the one-dimensional space of the likelihood as a function of the prior mass X.

The Nested Sampling procedure is then (Skilling, 2006)

- Start with N points $\Theta_1, \ldots, \Theta_N$ drawn from prior
  initialise $Z = 0$, $X_0 = 1$.
- Repeat for $j = 1, 2, \ldots, K$
  record the lowest of the current likelihood values as $\mathcal{L}_j$,
  set $X_j = \exp(-j/N)$
  set $w_j = X_{j-1} - X_j$
  increment Z by $\mathcal{L}_j w_j$
  replace point of lowest likelihood by new one drawn
  from prior subject to $\mathcal{L} > \mathcal{L}_j$
- Increment Z by $[\mathcal{L}(\Theta_1) + \ldots + \mathcal{L}(\Theta_N)]\ X_j/N$

where the last step refines the evidence integral with the information from the surviving N particles, with weight $N^{-1}X_j$. The final number of terms in the evidence summation is $N + K$.

* ◯ *

From these last paragraphs, calculating Z seems to actually be quite simple. But I skipped over a few important details. How to generate new particles from the prior, subject to the hard likelihood constraint? How to assign X values? How many particles should we consider (the value of N)? For how many steps K should the algorithm run to achieve convergence? Once these questions are answered we have an algorithm to calculate Z.

Understanding Nested Sampling

The key to the Nested Sampling algorithm lies in the mapping between the parameter space and the X-space, depicted in Fig. A.1, and once this is understood, most other aspects of the algorithm are simply implementation details. Let us see the reasoning behind it.

Since the prior is a probability distribution, it integrates to one over the entire parameter space:

$$1 = \int \pi(\Theta)\, d\Theta. \tag{A.3}$$

We define the prior mass associated with likelihoods greater than λ:

$$X(\lambda) = \int_{\mathcal{L}(\Theta)>\lambda} \pi(\Theta)\, d\Theta. \tag{A.4}$$

By construction, this is a decreasing function with

Likelihoods are strictly positive.

$$X(0) = 1 \qquad X(\lambda_{\max}) = X(\lambda = \mathcal{L}_{\max}) = 0. \tag{A.5}$$

As λ increases from 0 to $\lambda_{\max}$, the constrained likelihood function ($\mathcal{L}(\Theta) > \lambda$) occupies a smaller volume of the prior and the prior mass on that volume decreases.

Denoting the inverse function by $L(X)$, such that, $L(X(\lambda)) = \lambda$, this is a positive and decreasing function of X with

$$L(0) = \mathcal{L}_{\max} \qquad L(1) = 0 \tag{A.6}$$

and the evidence becomes a one-dimensional integral over $[0, 1]$:

$$Z = \int_0^1 L(X)\, dX\,. \tag{A.7}$$

Note that there is no trick in going from Eq. (A.2) to Eq. (A.7) and, as long as we can construct the function $L(X)$, we can simply integrate it (numerically) to obtain Z. If we could assign X values (say uniformly between 0 and 1) and calculate $L(X)$ for each, evaluating Z would indeed be simple. But we cannot. We are only able to go in the opposite direction, from values of L to values of X.

In a pool of X values, we know that X_i lies below X_{i-1} if and only if $L(X_i) > L(X_{i-1})$. Therefore, given a value $\mathcal{L}(\Theta_{i-1})$ with its associated X_{i-1}, if we sample uniformly from the prior subject to $\mathcal{L}(\Theta) > \mathcal{L}(\Theta_{i-1})$, we know for sure that the new X_i is uniformly distributed in $(0, X_{i-1})$. Sampling from the unconstrained prior gives us a definite starting point: $\mathcal{L}(\Theta_0)$ and $X_0 \sim U(0,1)$; from here we just progress with this statistical assignment of X values, until we are confident that the function $L(X)$ is well mapped.

The Nested Sampling algorithm uses a set N points initially sampled uniformly from the prior, instead of just one. But the reasoning is the same. The likelihood values for those N points are not necessarily sorted, but we can always assign, probabilistically, an X value to the lowest likelihood value and continue from there, sampling from the constrained prior distribution.

Diffusive Nested Sampling

The main challenge with Nested Sampling is to generate the new particles from the prior subject to the hard likelihood constraint. These new particles should also be independent of all the other ones currently in the population.

The simplest approach, as suggested by Sivia and Skilling (2006), is to copy one of the surviving $N-1$ points, evolve it with an MCMC targeting the prior distribution, and reject proposals that do not respect the likelihood constraint. Running this MCMC for long enough will produce a new independent particle that indeed follows the constrained distribution. But this method is increasingly inefficient, especially in the later stages, when the algorithm is exploring very constrained distributions.

More sophisticated sampling methods were also proposed, e.g. in the `Multinest` package by Feroz and Hobson (2008).[2] Indeed this approach has been very successful, and even saw some applications to the exoplanet field (Feroz et al., 2011a,b; Feroz and Hobson, 2014).

[2] See also Feroz et al. (2009, 2013).

However, as Feroz et al. (2011b) already noted, the `Multinest` package in particular and Nested Sampling in general, tend to suffer from the curse of dimensionality, with sampling efficiency decreasing rapidly with the dimension of Θ. This is particularly problematic if the posterior is multimodal or highly correlated.

Recently, Brewer et al. (2011) introduced a new algorithm, which they called Diffusive Nested Sampling (DNS), designed to be as flexible and general as an MCMC, but also capable of efficiently exploring difficult constrained distributions. The algorithm introduces a slight but important improvement to the classic Nested Sampling approach, in that it attempts to sample from a *mixture* of successively constrained distributions, instead of using one single hard constraint at each step.

DNS starts by generating a particle from the prior (call this distribution $p_{\mathcal{L}_0}$, with $\mathcal{L}_0 = 0$) and evolving it with an MCMC, storing all the intermediate likelihood values in an accumulated array. After a

given number of iterations, it finds the $1 - e^{-1} \sim 63\%$ quantile of all the likelihood values, and records it as $\mathcal{L}_1$; this creates a new *level* occupying about e^{-1} times the mass of $p_{\mathcal{L}_0}$ (in this case, the prior mass). All the likelihood values lower than $\mathcal{L}_1$ are then removed from the accumulated likelihood array.

At this point, Nested Sampling would continue sampling from the prior constrained to $\mathcal{L}_1$ (call it $p_{\mathcal{L}_1}$), via MCMC for example. In contrast, DNS attempts to sample from a weighted sum of the two distributions $p_{\mathcal{L}_0}$ and $p_{\mathcal{L}_1}$. An MCMC is used to evolve the particle subject to this mixture of distributions, and once enough samples have been obtained from $p_{\mathcal{L}_1}$, we again find the $1 - e^{-1}$ quantile of all the likelihood values, and record it as $\mathcal{L}_2$. Likelihood values smaller than $\mathcal{L}_2$ are removed. The particle then explores a mixture of $p_{\mathcal{L}_0}$, $p_{\mathcal{L}_1}$ and $p_{\mathcal{L}_2}$ and this process continues until a maximum number of levels is created. Once all the levels have been obtained, the particle simply continues to explore the mixture of all the levels until the algorithm is terminated.

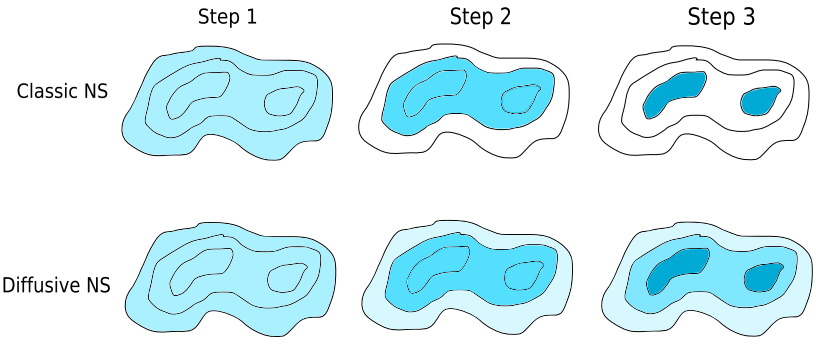


Figure A.2: Illustration of the constrained distributions that Classic Nested Sampling and Diffusive Nested Sampling must sample as they progress. In the classic scheme, the bimodal distribution of step 3 will be quite difficult to sample from. Exploring instead the mixture distribution makes travelling between isolated modes more likely. Adapted from Brewer et al. (2011).

Each time a new level is created, its constrained distribution covers about e^{-1} times as much prior mass as the last distribution. Therefore, the X-value of the kth level can be estimated as $\exp(-k)$, an estimate which will be refined later on.

In order to create the mixture of distributions we need to provide a weighting scheme for each component. Simple uniform weights for all distributions would work, albeit inefficiently. Brewer et al. (2011) proposed exponentially-decaying weights with a scale length Λ, which describes how far (down in likelihood) the particle can go in order to explore more freely.

When the desired number of levels has been created, the weights can be changed to uniform, and further samples are drawn from all the component distributions. The algorithm can then continue to sample for as long as required, with the evidence and posterior samples converging to their true values.

As the levels are being created, their actual X-values can be modified from the theoretical expectation $X_{j+1} = e^{-1}X_j$. This means the weight of each distribution is actually different and the exploration is thus not completely correct. The X-values can nevertheless be corrected. At a given level j, the values of the likelihood will be higher

than the upper level's likelihood cut-off a fraction X_{j+1}/X_j of the time. Thus, we can use the actual fraction of samples in which this happens as an estimate of the true ratio of the X-values for consecutive levels.

* ◯ *

The DNS algorithm is essentially an application of the Metropolis-Hastings algorithm to a distribution other than the posterior. Changing the target distribution improves upon other MCMC algorithms by providing the value of the evidence in one single run and being less sensitive to the presence of complicated features in the posterior. Classic Nested Sampling also shares these advantages. But DNS improves upon the classic algorithm by alleviating the problem of sampling from the constrained likelihood function. Because the target distribution used by DNS always includes the prior distribution as part of the mixture, sampling from posteriors with substantial multimodality is still possible and even efficient.

Finite mixture models

DNS can be applied quite generally to many sampling problems. For our purposes, we would like to apply this algorithm to the trans-dimensional problem that we studied in Chapter 2. This means that, in particular, we need to sample the posterior for the N_p parameter and need to change the dimensionality of the parameter space while the algorithm runs. The application of DNS to such problems has been developed by Brewer (2014), and I describe it briefly here.

In a practical implementation, and because of the way Nested Sampling and DNS work, we only need to define a set of proposal distributions that will allow us to sample from the prior. There are a few theoretical constraints on these proposals but they can still be quite general distributions. With the proposals defined, the DNS algorithm can simply incorporate the hard likelihood constraints and sample from the mixture of distributions.

In Brewer (2014), a set of proposal distributions are presented for three different classes of parameters denoted N, $\{\boldsymbol{x}_i\}$ and $\boldsymbol{\alpha}$. These correspond to the number of components, each component's individual parameters, and a set of hyperparameters. In the notation used in this Thesis, these have direct correspondence to N_p, $\boldsymbol{\theta}$, and $\boldsymbol{\alpha}$, where N_p is the number of Keplerians, $\boldsymbol{\theta}$ the orbital parameters of each, and $\boldsymbol{\alpha}$ the extra parameters that control for the semi-amplitudes and orbital periods.

Being arguably more interesting that the others, I will discuss here only the proposal distribution for N_p. Changing the value of N_p during the run of the algorithm changes the dimension of the parameter space, and therefore must also change the set of parameters $\boldsymbol{\theta}$ (but not necessarily $\boldsymbol{\alpha}$). These proposals are often called birth/death proposals (see for example Stephens, 2000).

See Brewer (2014) for the complete description of the other proposals.

The prior distribution for N_p is uniform between 0 and $N_{p,\max}$. A

new value N_p' is calculated following

$$N_p' = \mathrm{mod}(N_p + \delta N; N_{p,\max} + 1) \tag{A.8}$$

where δN is drawn from the heavy tailed distribution

$$\mathrm{mod}(u + 10^{1.5-6a} b; 1) \quad \text{with } u \sim U(0,1) \\ a \sim U(0,1) \text{ and } b \sim N(0,1) \tag{A.9}$$

This proposal is such that δN takes values ± 1 most of the time, but values of order $N_{p,\max}$ are also possible, allowing for fast exploration of distributions as wide as the prior.

When $N_p' > N_p$, we are proposing to add Keplerians to the model, and the parameters of these new Keplerians are drawn from their priors conditional on the current value of the hyperparameters. Otherwise, $N_p - N_p'$ of the current N_p Keplerians are selected randomly for removal, with equal probability.

Details on implementation

The DNS algorithm has been implemented in the `DNest4` package, as described in Brewer and Foreman-Mackey (2016). `DNest4` is an open source, multi-threaded implementation, which also provides a template for finite mixture models (Brewer, 2014).

This software has been used to sample the posterior distribution in an exoplanet detection setting (Brewer and Donovan, 2015), and I have extended it with a complete Keplerian solver and the introduction of Gaussian Processes (Faria et al., 2016a). After setting the probabilistic model, in the form of the prior distributions and the likelihood, the software is of remarkably general application.

A few numerical parameters have to be defined as options to the DNS algorithm, the most important ones being:

- `NewLevel`
  a new level will be created once `NewLevel` MCMC steps have resulted in `NewLevel` likelihood values above the current top level.

- `MaxLevels`
  Sets the maximum number of levels. Very useful values for this parameter are 0 (zero) which tells `DNest4` to automatically determine the maximum number of levels and 1 which makes the algorithm explore the prior distribution only. For typical datasets, this parameter ranges from 30 to about 200.

- Λ, the backtracking scale length;
  Controls the degree to which particles are allowed to backtrack down, while the levels are being built.

The outputs of a `DNest4` run consist of three files: one with the samples of parameter values that represent the mixture of constrained priors (not the posterior distribution), one with the information about all the levels created by the algorithm, such as their likelihoods and

values of $\log(X)$, and a file with metadata about each sample (such as which level it belongs to).

These outputs require further post-processing, in order to calculate the value of the evidence and obtain posterior samples. A number of diagnostic plots are also produced which can be used to monitor and check the progress of a run. Convergence of the algorithm is tested heuristically by analysing these plots.

In the listing below, I show the RV model as a higher level implementation to use in the `DNest4` package. In the interest of reproducible research, the complete working version of the code is available in a new open source package called `kima`, at the following address: https://github.com/j-faria/kima. The version shown here does not include the GP component to model stellar activity RV variations, since that was already published with Faria et al. (2016a).

```
RVmodel::RVmodel()
:objects(5, Np,max , false, ConditionalPrior())
,mu(N)
,C(N, N)
{}

void RVmodel::from_prior(RNG& rng)
{
  //<-- here should sample orbital parameters from respective priors *//

  vsys = min(vobs) + (max(vobs) - min(vobs)) * rng.rand();

  #if obs_after_fibers
    // between 0 m/s and 30 m/s
    fiber_offset = 30e-3*rng.rand();
  #endif

  // cauchy centered at 0.001 (data in km/s)
  s = exp(-6.908 + tan(M_PI*(0.97*rng.rand() - 0.485)));

  #if tstudent_likelihood
    nu = exp(log(0.1) + log(1000.)*rng.rand());
  #endif

  calculate_mu();
}

void RVmodel::calculate_mu()
{
  mu = vsys

  double P, A, phi, ecc, w, f, v, ti;
  for(j in Np)
  {
    for(i in N)
    {
      f = true_anomaly(t[i], P_j, ecc_j, t[0]-(P_j*phi_j)/(2.*M_PI));
      v = K_j*(cos(f+w_j) + ecc_j*cos(w_j));
      mu[i] += v;
    }
  }

  #if obs_after_fibers
    for(i in N)
      if (i>=index_fibers) mu[i] += fiber_offset;
```

```
  #endif
47

49 double RVmodel::log_likelihood() const
{
51  #if tstudent_likelihood
    double logL = 0.;
53  double var;
    for(i in N)
55  {
      var = sigma[i]*sigma[i] + s * s;
57    logL += lngamma(0.5*(nu + 1.)) - lngamma(0.5*nu)
          - 0.5*log(M_PI*nu) - 0.5*log(var)
59        - 0.5*(nu + 1.)*log(1. + pow(y[i] - mu[i], 2)/var/nu);
    }
61  #else
    double halflog2pi = 0.5*log(2.*M_PI);
63  double logL = 0.;
    double var;
65  for(i in N)
    {
67    var = sigma[i]*sigma[i] + s * s;
      logL += - halflog2pi - 0.5*log(var)
69        - 0.5*(pow(y[i] - mu[i], 2)/var);
    }
71  #endif

73  return logL;
}
```

Listing A.1: Implementation of the RV model to use with `DNest4`.

Appendix B: *OPEN*

Over the four years of work that culminated in this Thesis, I have analysed hundreds of different RV datasets. This means that some common tasks had to be done efficiently and reproduced many times. With the motivation of making this easier, and also to concentrate the methods I developed in a single package, I created a computational platform called `OPEN`. This brief appendix will describe `OPEN`'s top-level implementation and some of its features.

`OPEN`'s name is related to it being open-source but is also a recursive acronym for the **O**pen **P**latform for **E**xoplanet a**N**alysis.

Written mainly in `Python`, but with specific modules in other languages, `OPEN` builds on the system of '`magic`' commands of the `IPython` package (Pérez and Granger, 2007). It provides a console with custom commands geared towards the analysis of RV data, while retaining the capabilities of the full `IPython` environment.

Running the file `open.py` starts the interactive console:

```
Welcome to OPEN v0.0.5
Created by João Faria | joao.faria@astro.up.pt

Type "listcommands" for a list of the available commands
For help on a specific command, type "command -h" or see http://j-faria.github.io/OPEN
OPEN [1]:
```

after which the user can start to input commands, as in a typical `IPython` console. As stated in the welcome text, a list of available commands can be accessed by typing `listcommands` and to view any command's help text we can type its name followed by `-h`.

The main component of the platform is a class `RVseries`, which holds the RV observations together with a number of methods. Other functions can act on instances of this class. To read a file with RV observations of HD 56274, for example, we could type

```
OPEN [1]: read -h
Usage:
    read <file>...
    read <file>... [-d] [--skip=<sn>] [-v] [--quiet] [--nomps]
    read -h | --help
Options:
    -d                  Set this as default system.
    -v --verbose        Verbose output about data just7 read.
    --quiet             Do not print any output.
    --skip=<sn>         How many header lines to skip [default: 2].
    --nomps             Do not convert data to m/s
    -h --help           Show this help message.


OPEN [2]: read HD56274_harps.rdb -d --nomps
```

By default, OPEN easily reads the files produced by the HARPS DRS, with data organized in columns and a 2-line header. The option `-d` set this as the default system, meaning that a variable called `default`, an instance of the class `RVseries`, is now available:

```
OPEN [3]: default
  Out<3>: System with RV info from {'HD56274_harps.rdb': [235, 0]}
```

This variable holds the data that was read from the file. The attributes `default.time`, `default.vrad` and `default.error`, for example, hold the times of observation, the RV measurements and their uncertainties. Other variables are stored in the `default.extras` attribute. The `-d` setting is optional so that the user can keep different systems in different variables, as in

```
OPEN [4]: system1 = %read HD56274_harps.rdb --nomps
```

Note that the `%` character may be needed when OPEN's command is not the first thing typed on a given line.

After reading the data, we can show various quantities with the `plot` command:

```
OPEN [5]: plot obs
OPEN [6]: plot rhk
```

which create the figures below with the timeseries of RVs and log $R'_{\rm HK}$ measurements. These and other variables could also be visualised together using `plot obs --together=rhk`.

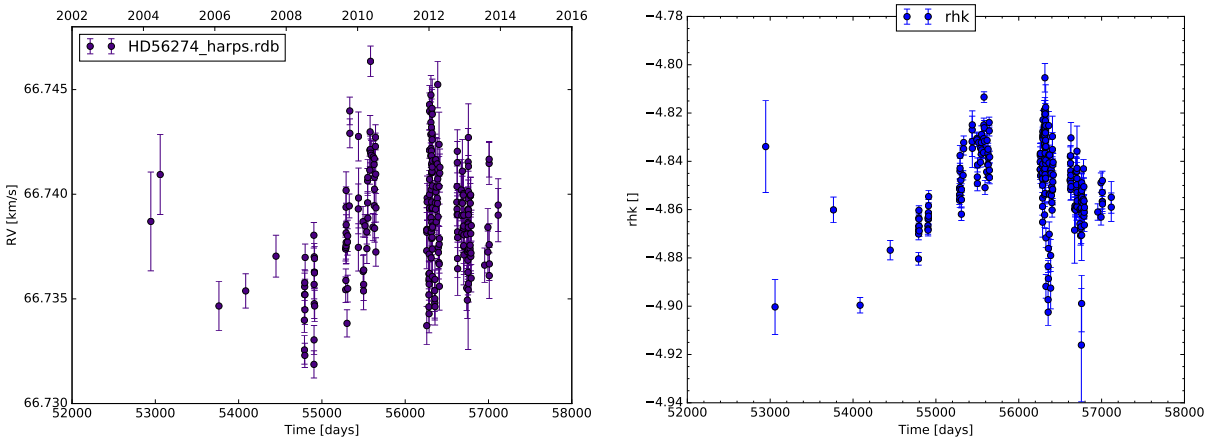


Figure A.3: Plots of the RV and log $R'_{\rm HK}$ timeseries for HD 56274 created by OPEN.

Also part of the core of OPEN are the implementations of various algorithms to calculate different periodograms, all accessible with the `per` command. Currently, there are implementations of the Generalized Lomb-Scargle (GLS, Zechmeister and Kürster 2009), Bayesian Generalized Lomb-Scargle (BGLS, Mortier et al. 2015), Lomb-Scargle (LS, Lomb 1976; Scargle 1982), Hoeffding-test (Zucker, 2015) and multiband (VanderPlas and Ivezić, 2015) periodograms.

The distinct versions of the Lomb-Scargle periodogram differ in whether they consider the observational uncertainties, a free constant

offset or a Bayesian generalisation of the method. All periodograms are also different in their computational efficiency, and thus how they scale with the number of observations.

A number of options can be given to `per` to increase the frequency resolution, decrease the lowest frequency at which to calculate the periodogram (by default, the timespan of the data) and calculate False Alarm Probabilities (FAP) based either on theoretical distributions or Monte Carlo bootstrap simulations.

See `per -h`

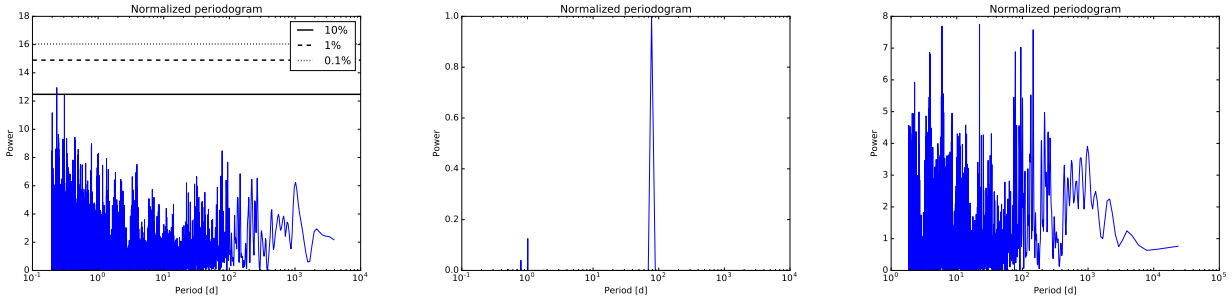


Figure A.4: Different periodograms of the RVs of HD 126793. From left to right: GLS (option `-g`), BGLS (option `-m`) and LS (option `-l`). The FAPs in the first plot were calculated with the option `--bfap`.

Examples of these periodograms are shown in the figure above, using the RV data of HD 126793. They were calculated using the command `per obs` with different options: `-g` (default), `-m`, `-l`. Note that the FAP lines in the GLS periodogram were calculated using Monte Carlo simulations (option `--bfap`) which is quite computationally demanding.

* ◯ *

Many commands in `OPEN` are simply utility functions which perform simple actions when applied to or using the data from `RVseries` instances. As examples:

```
OPEN [7]: mod d2    # set the model as a parabola
OPEN [8]: fit       # fit the model to the RV observations and subtract it
                    # coefficients are saved in default.model

OPEN [9]: rotation  # estimate rotation period from activity-rotation relations
                    # using the current system's logR'hk measurements
                    # optionally ask for B and V magnitudes

OPEN [10]: create    # create simulated RV curves

OPEN [11]: restrict  # Select specific data points,
                     # depending on their uncertainty, timestamp, or within
                     # some year interval, for example.
                     # see also the --gui option.

OPEN [11]: saverdb   # save the current system's data as an rdb file
```

One important goal in the development of OPEN was extensibility. The objective is that the user is not tied to a given syntax and is able to develop their own code more easily. This led to the possibility of extending OPEN with *macros* or *plugins*.

Macros are Python scripts that use some of OPEN's commands in order to manipulate different kinds of data. To run these scripts from within the OPEN console, the user can type

```
OPEN [8]: run -i /path/to/macro.py
```

which makes available to the script a number of variables already defined. Within the script, code such as

```
O = get_ipython().magic
```

allows the use of OPEN's commands as

```
system = O('read %s --nomps' % data_file)
O('restrict --index=71 --noask -n system')
```

Many of the figures, tables, and indeed most of the results presented in this Thesis were created with OPEN macros that combined these atomic operations on RVseries instances.

Another possibility to extend OPEN, even if less developed at the moment, is to write new commands in the form of plugins. If the user creates the directory structure OPEN/plugins/NewPlugin, and types

```
OPEN [8]: NewPlugin
```

OPEN will execute all the code in a script with name run.py, placed inside the directory OPEN/plugins/NewPlugin.

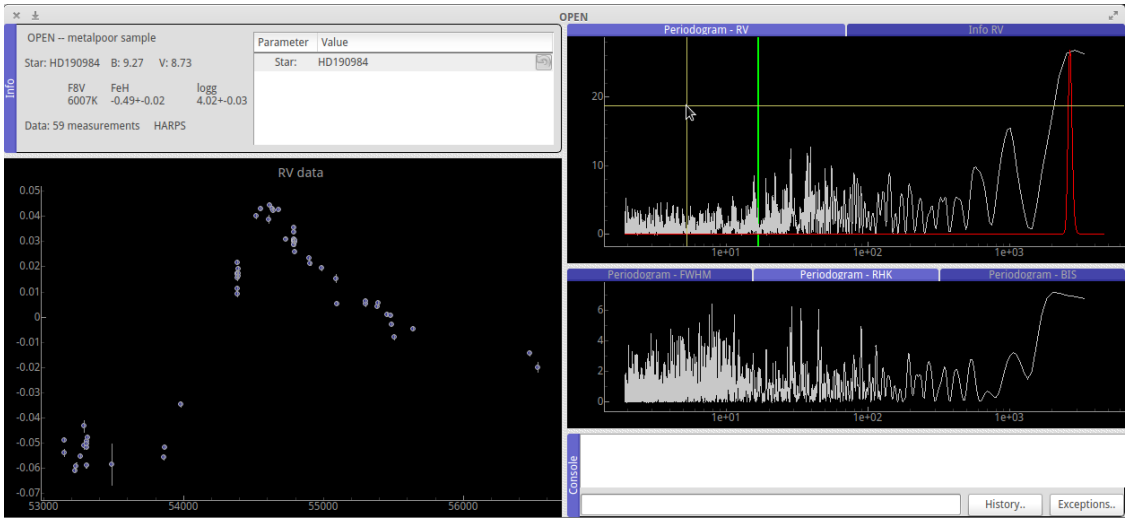


Figure A.5: Interface for the visualisation of data from the metal-poor sample, implemented as an OPEN plugin.

This plugin system has been used, for example, to create a graphical interface which proved helpful in the visualisation of data from the metal-poor large program:

```
OPEN [8]: metalpoor    # for demonstration, not available by default
```

which uses the `pyqtgraph` library[3] to render the RV timeseries and periodograms of some activity indicators (see figure above). Plugins are inherently user-defined and therefore are not suited to be included in the basic distribution of the project itself.

[3] See pyqtgraph.org.

`OPEN` is free software and completely open-source and has been developed by myself, and hosted on `GitHub`, since 2013. It can be accessed at github.com/j-faria/OPEN. The figure below shows the number of contributions to the project over the years, counting the commits made to the underlying source code repository.

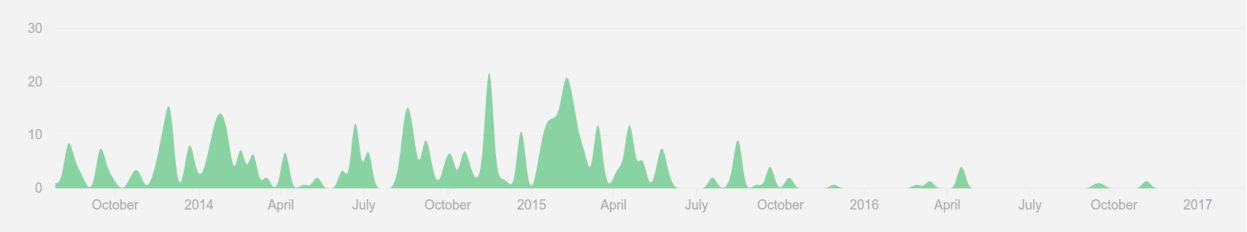


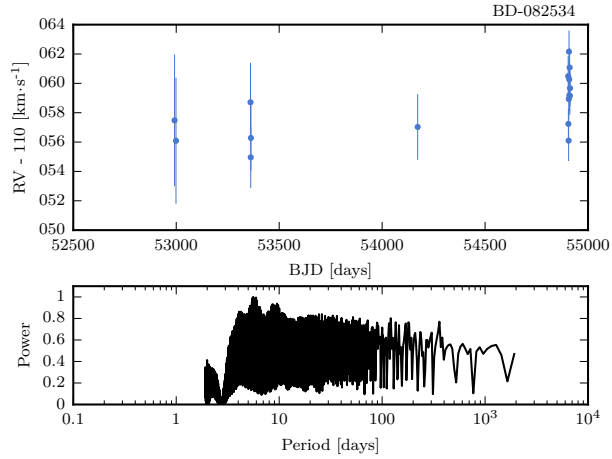
Figure A.6: Contributions to `OPEN`'s repository on `GitHub` over time.

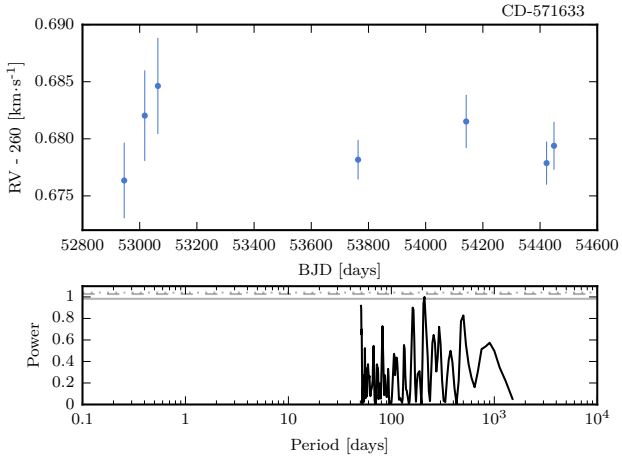
The project currently lacks documentation, tests and, most importantly, users. It nevertheless constitutes the computational backbone of this Thesis, and answers many of my goals for reproducibility. As a result, virtually all the figures in this Thesis were created using `OPEN` commands.

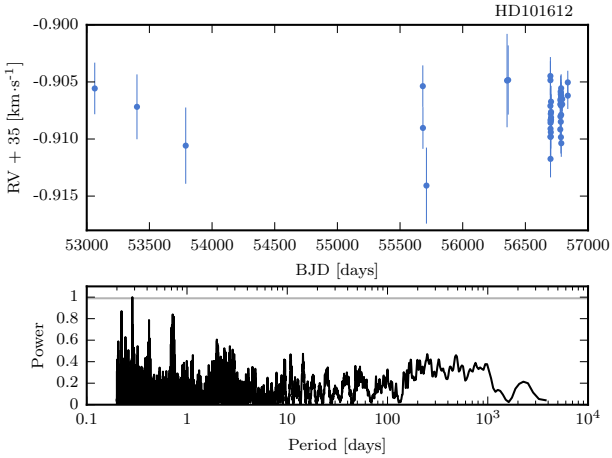
Appendix C:
Figures and Tables

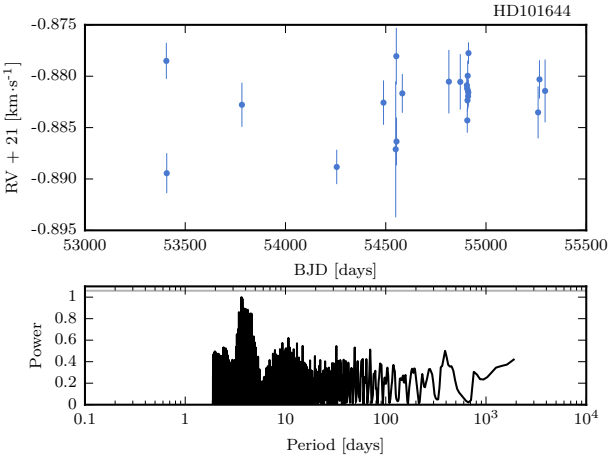
This Appendix presents a set of figures and tables to describe our metal-poor sample. First, the 109 RV timeseries and the corresponding GLS periodograms are shown. All the RVs are in km·s^{-1}, and in some cases an arbitrary offset was subtracted for visual clarity. Observations taken before and after the HARPS fiber change are color-coded. For most stars, the RVs obtained after the fiber change have been offset to match the weighted average of the remaining observations. Exceptions to this are the giant planet hosts. In the periodograms, the power has been normalized to 1 and the 10%, 1% and 0.1% FAP levels are also shown as horizontal solid, dashed and dotted lines, respectively. These have been calculated with the analytical formula given, e.g., by Zechmeister and Kürster (2009).

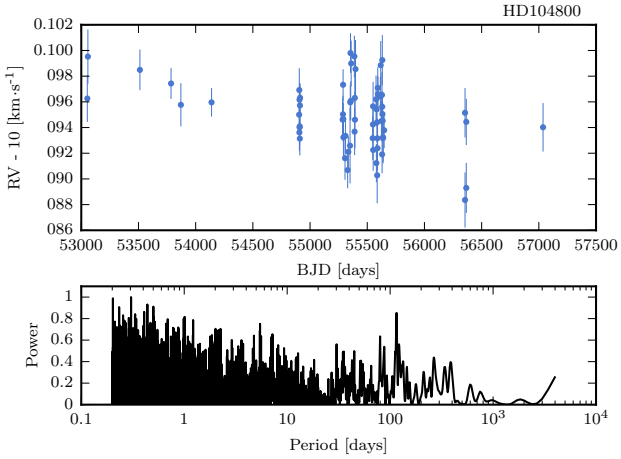
Note that in some figures these lines are not visible.

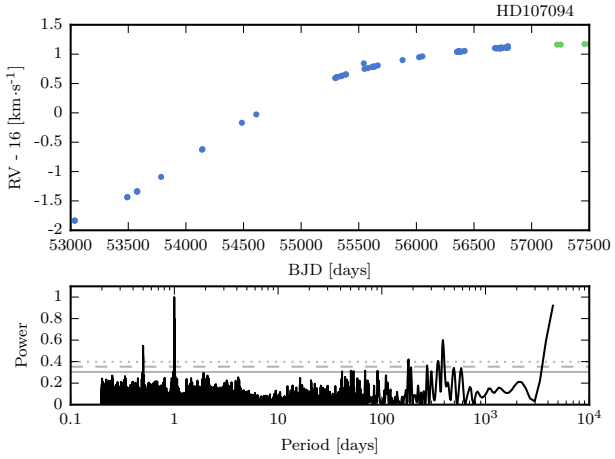


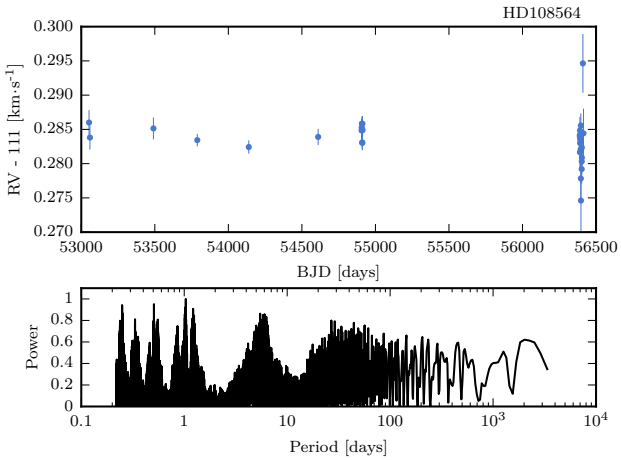


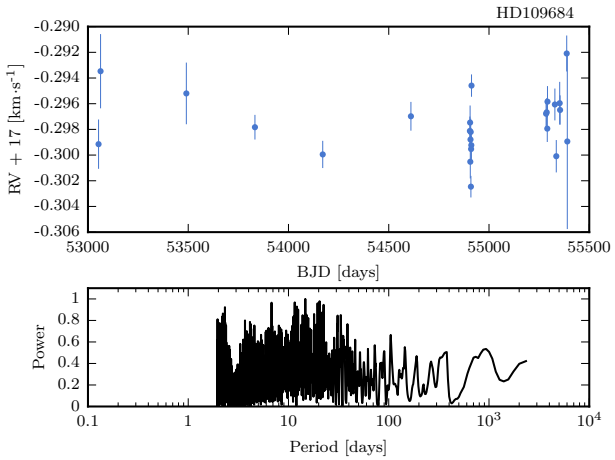


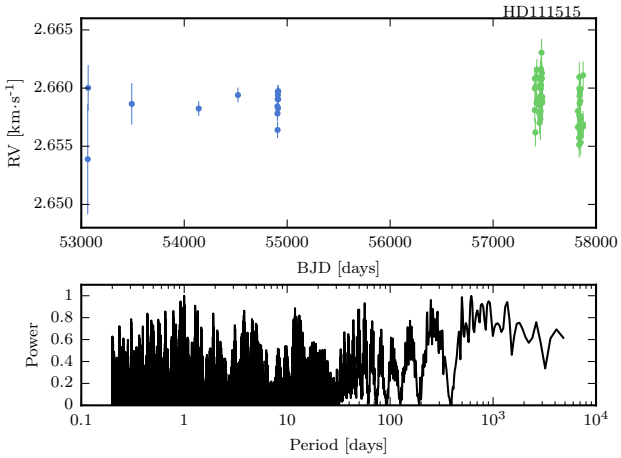


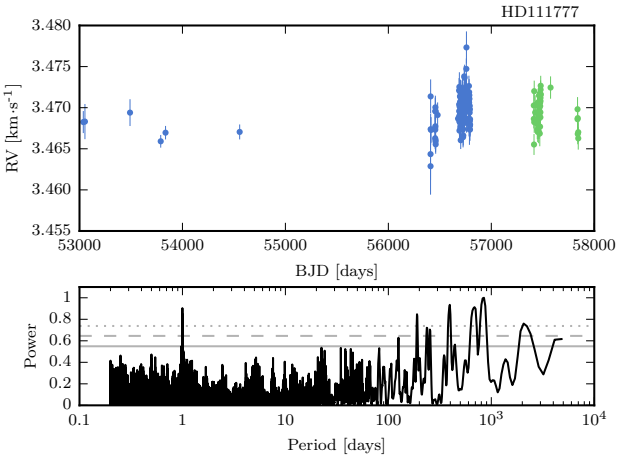

HD104800
RV - 10 [km·s⁻¹]
BJD [days]
Power
Period [days]

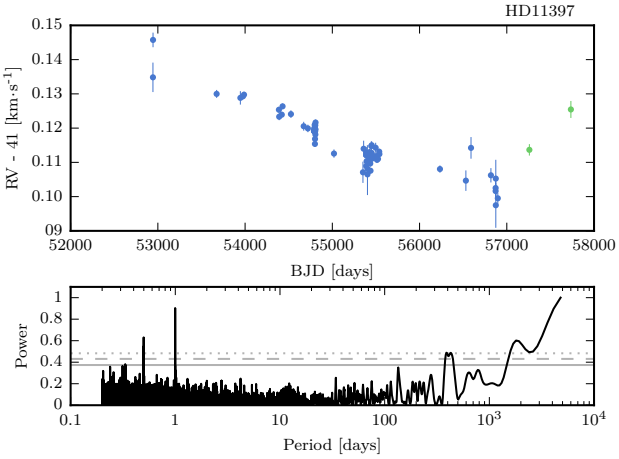

HD107094
RV - 16 [km·s⁻¹]
BJD [days]
Power
Period [days]

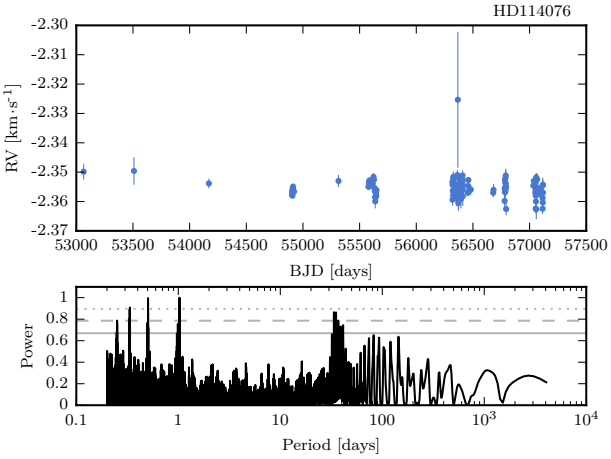

HD108564
RV - 111 [km·s⁻¹]
BJD [days]
Power
Period [days]

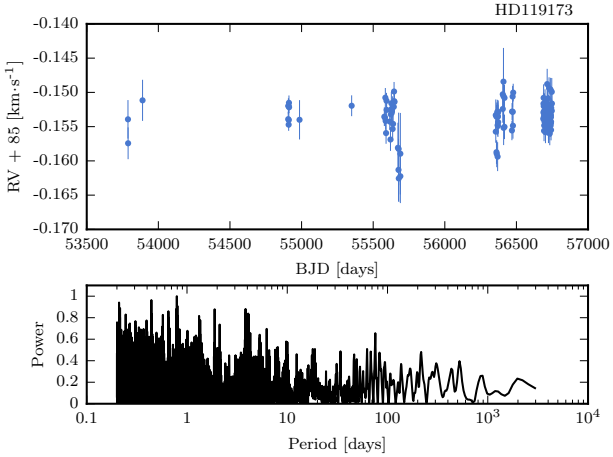

HD109684
RV + 17 [km·s⁻¹]
BJD [days]
Power
Period [days]

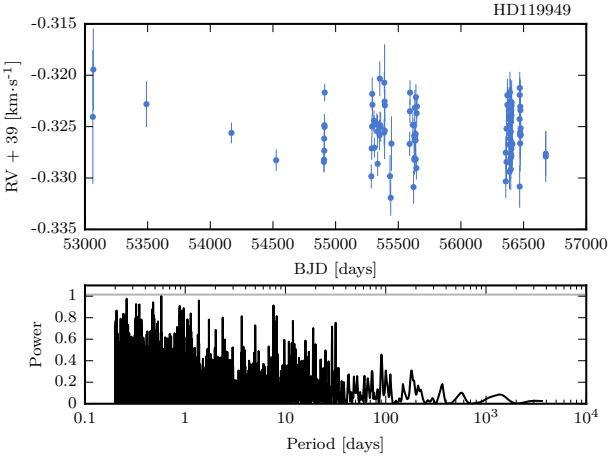

HD111515
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

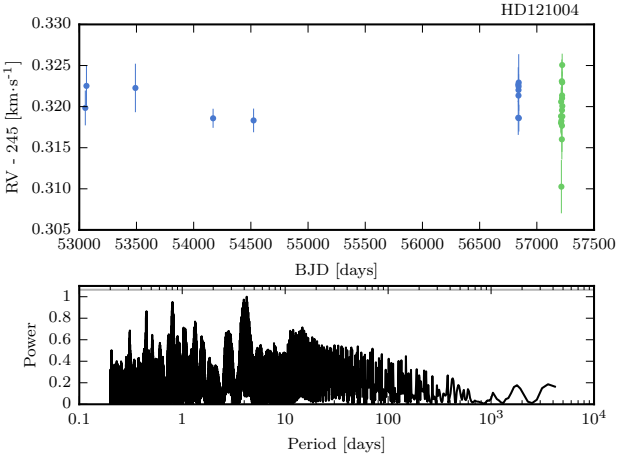

HD111777
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

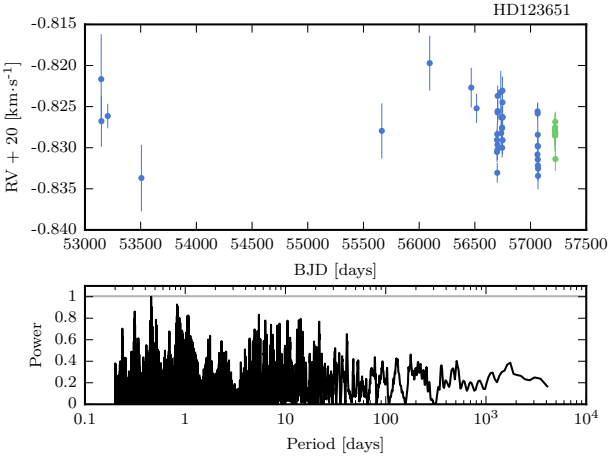

HD11397
RV - 41 [km·s^{-1}]
BJD [days]
Power
Period [days]

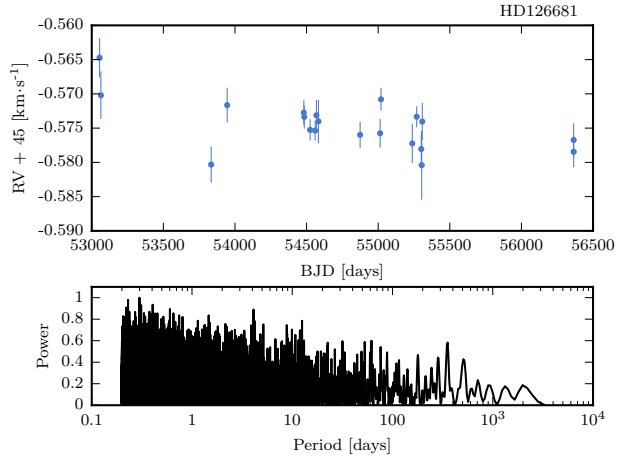

HD114076
RV [km·s^{-1}]
BJD [days]
Power
Period [days]

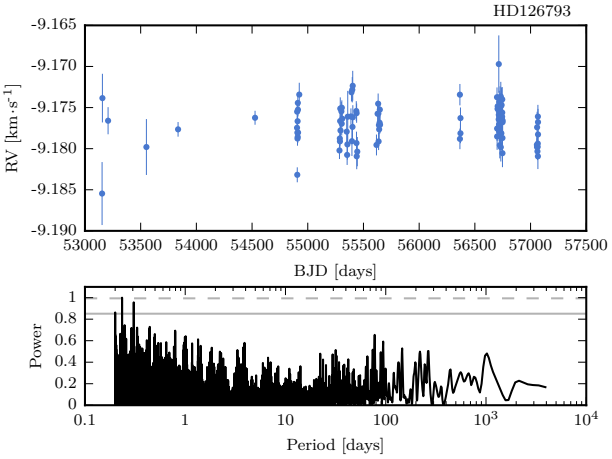

HD119173
RV + 85 [km·s^{-1}]
BJD [days]
Power
Period [days]

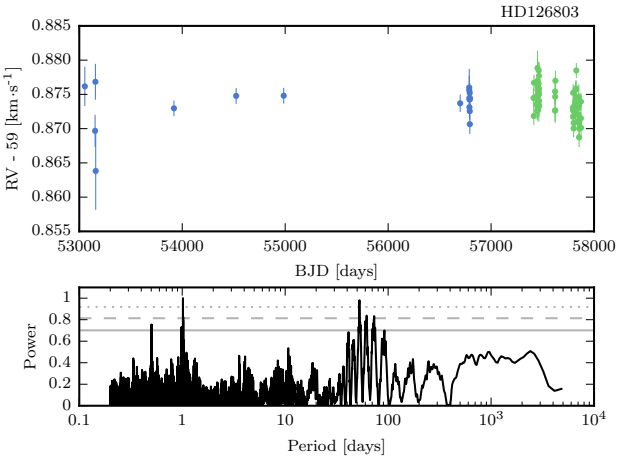

HD119949
RV + 39 [km·s^{-1}]
BJD [days]
Power
Period [days]

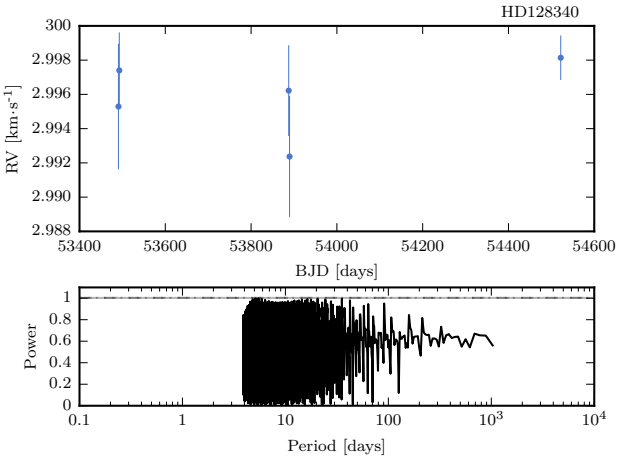

HD121004
RV - 245 [km·s^{-1}]
BJD [days]
Power
Period [days]

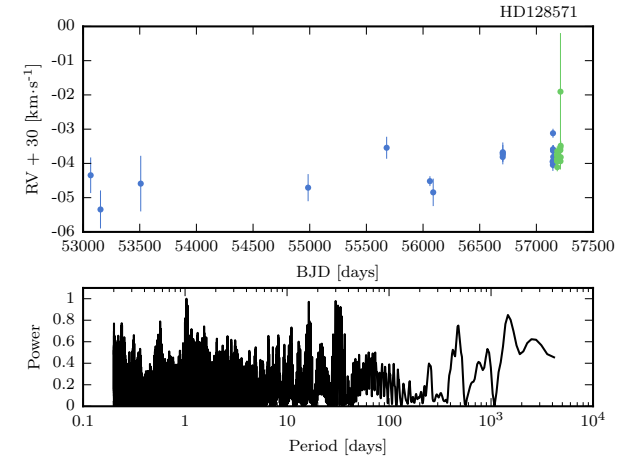

HD123651
RV + 20 [km·s^{-1}]
BJD [days]
Power
Period [days]

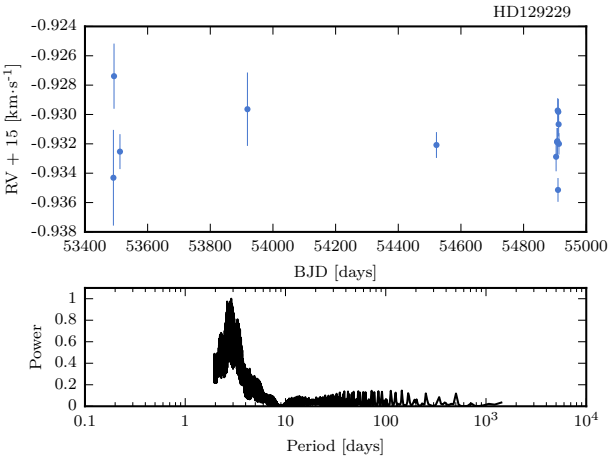

HD126681
RV + 45 [km·s⁻¹]
BJD [days]
Power
Period [days]

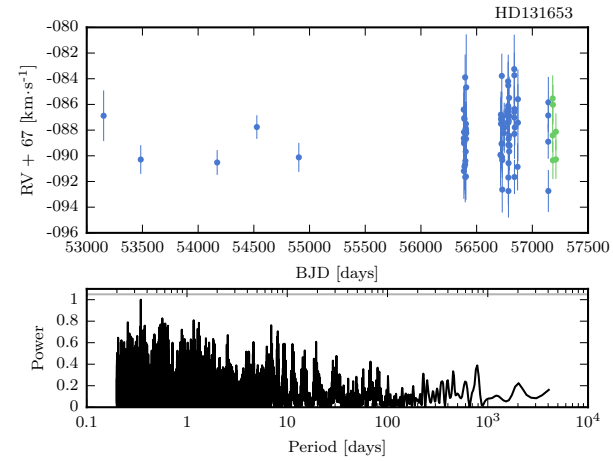

HD126793
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

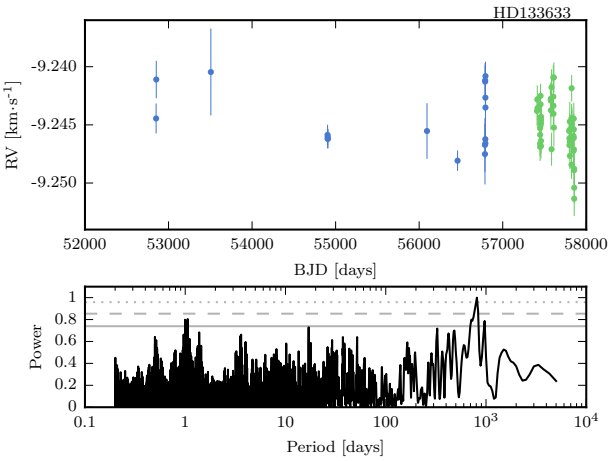

HD126803
RV - 59 [km·s⁻¹]
BJD [days]
Power
Period [days]

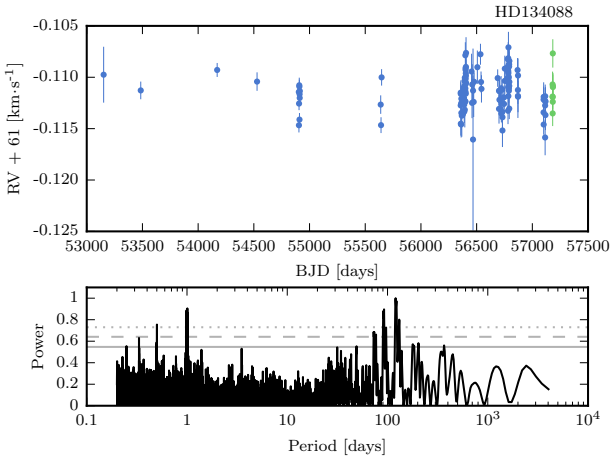

HD128340
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

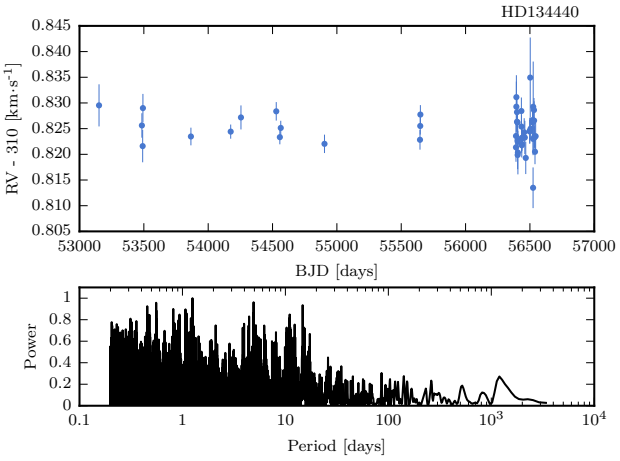

HD128571
RV + 30 [km·s⁻¹]
BJD [days]
Power
Period [days]

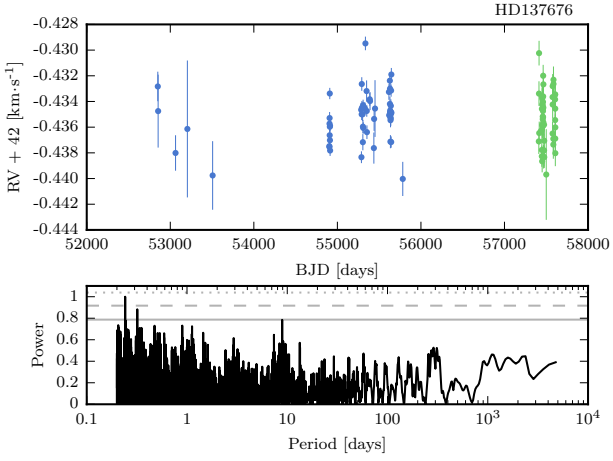

HD129229
RV + 15 [km·s⁻¹]
BJD [days]
Power
Period [days]

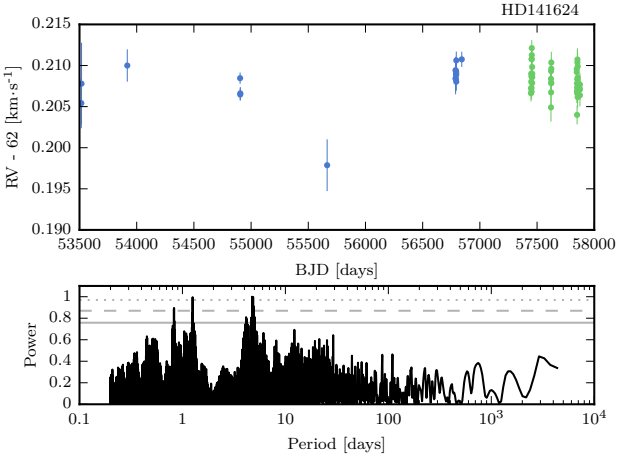

HD131653
RV + 67 [km·s⁻¹]
BJD [days]
Power
Period [days]

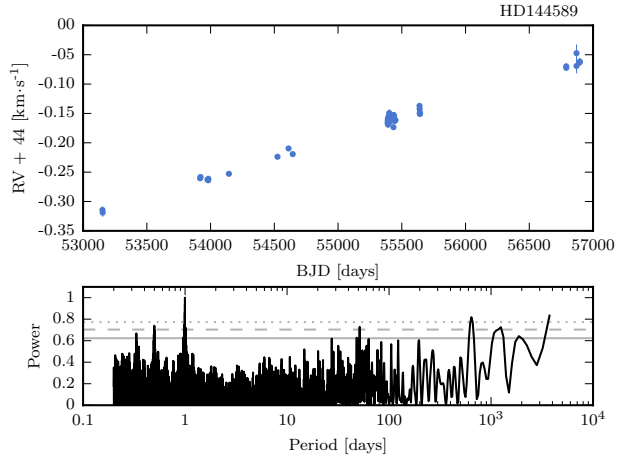

HD133633
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

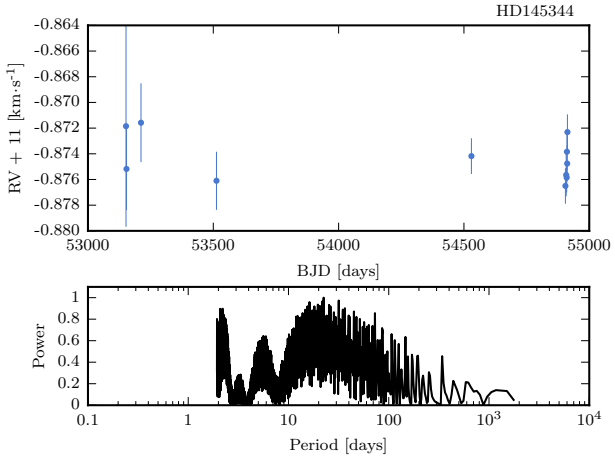

HD134088
RV + 61 [km·s⁻¹]
BJD [days]
Power
Period [days]

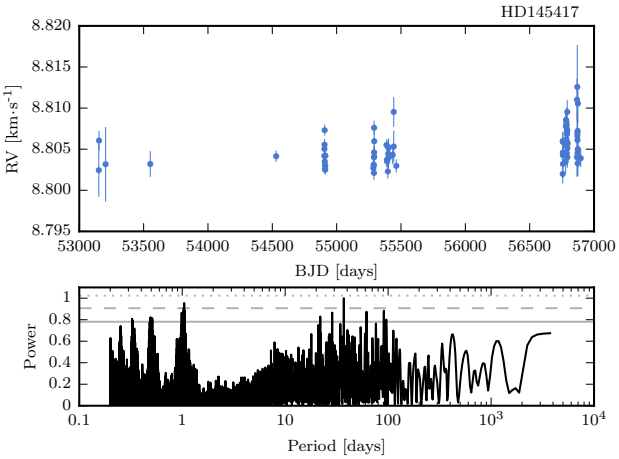

HD134440
RV - 310 [km·s⁻¹]
BJD [days]
Power
Period [days]

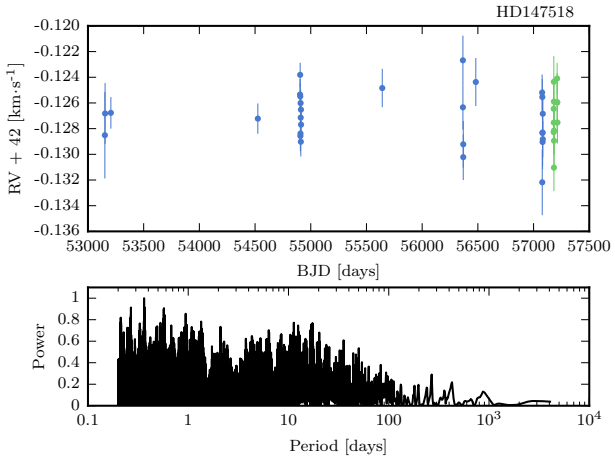

HD137676
RV + 42 [km·s⁻¹]
BJD [days]
Power
Period [days]

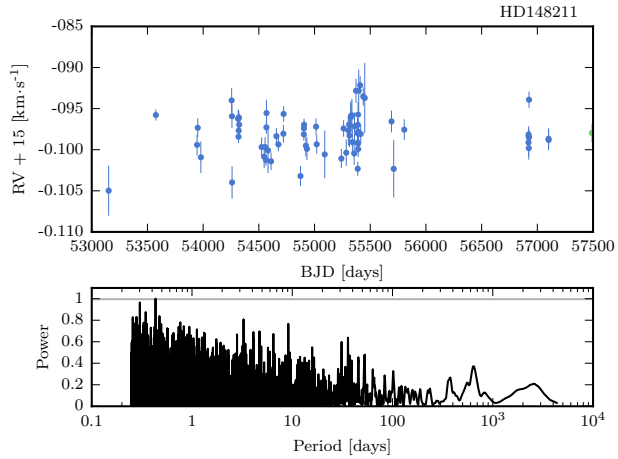

HD141624
RV - 62 [km·s⁻¹]
BJD [days]
Power
Period [days]

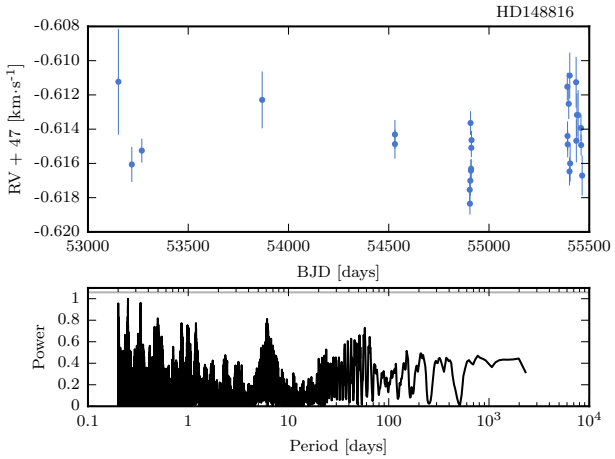

HD144589
RV + 44 [km·s⁻¹]
BJD [days]
Power
Period [days]

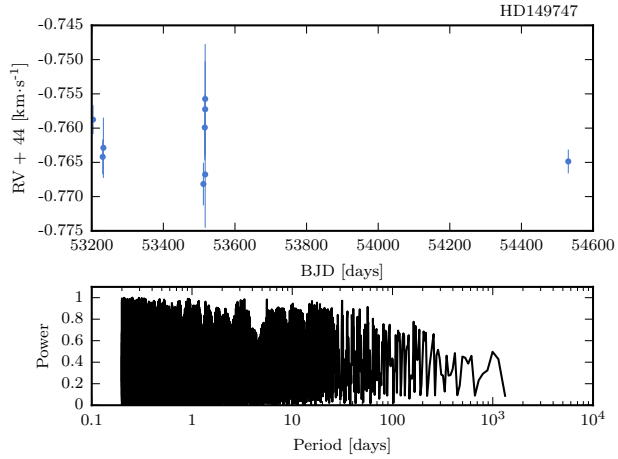

HD145344
RV + 11 [km·s⁻¹]
BJD [days]
Power
Period [days]

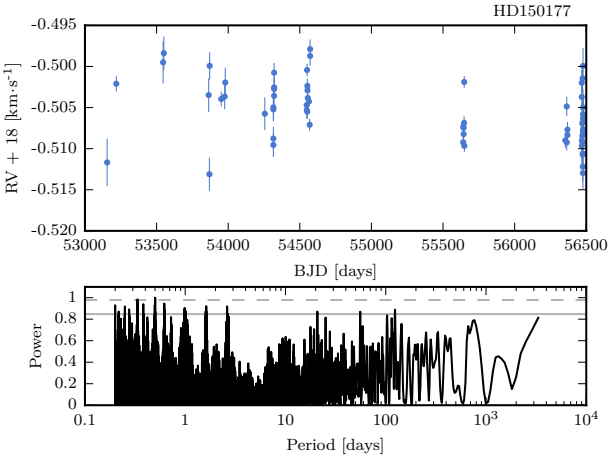

HD145417
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

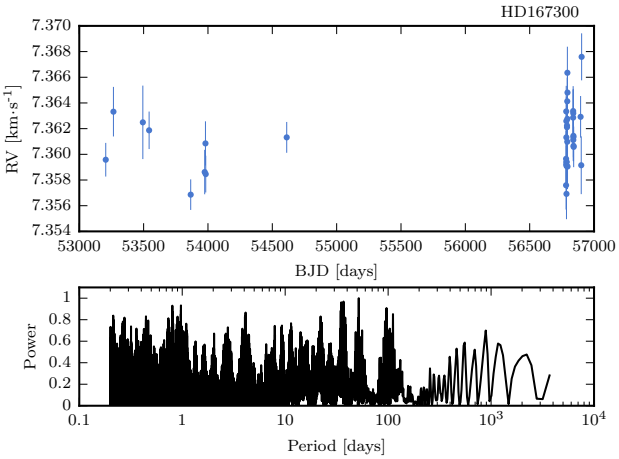

HD147518
RV + 42 [km·s⁻¹]
BJD [days]
Power
Period [days]

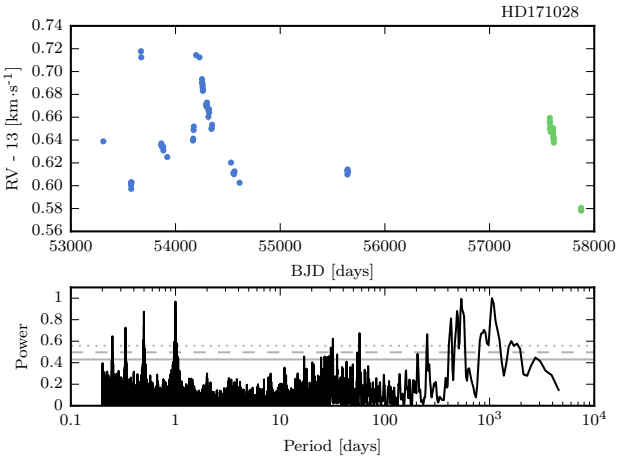

HD148211
RV + 15 [km·s⁻¹]
BJD [days]
Power
Period [days]

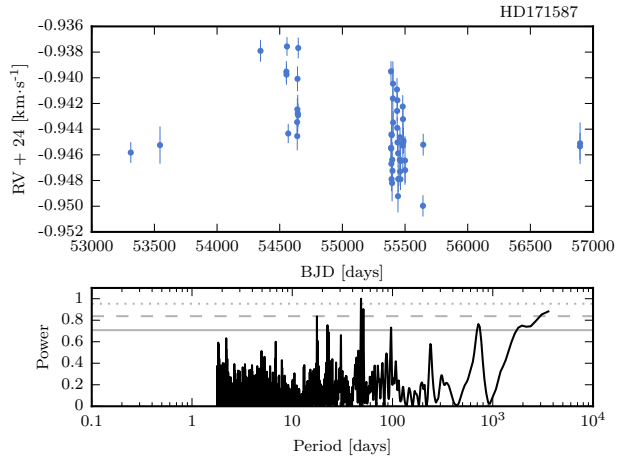

HD148816
RV + 47 [km·s⁻¹]
BJD [days]
Power
Period [days]

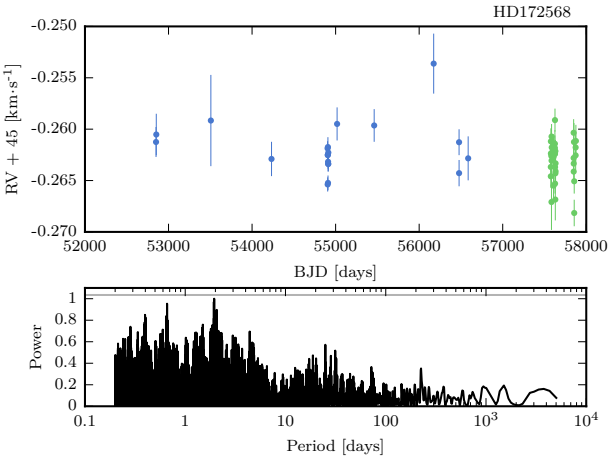

HD149747
RV + 44 [km·s⁻¹]
BJD [days]
Power
Period [days]

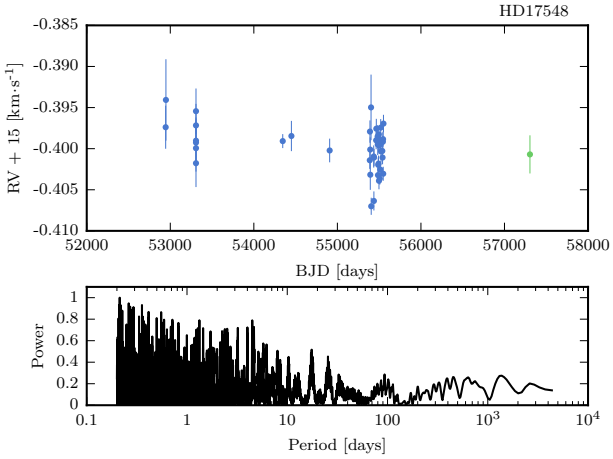

HD150177
RV + 18 [km·s⁻¹]
BJD [days]
Power
Period [days]

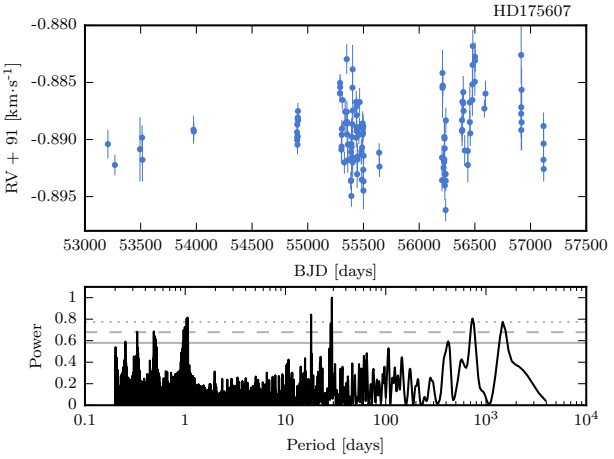

HD167300
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

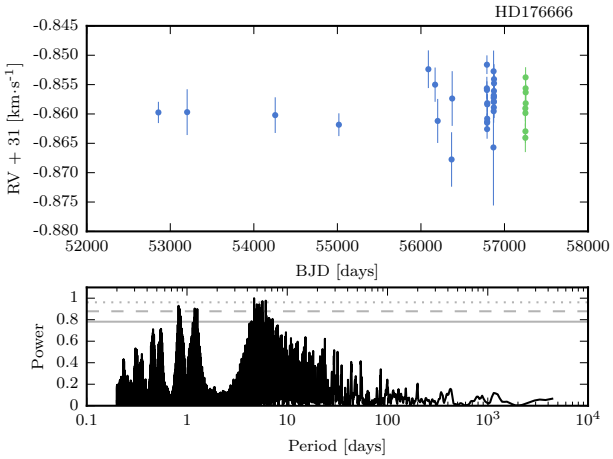

HD171028
RV - 13 [km·s⁻¹]
BJD [days]
Power
Period [days]

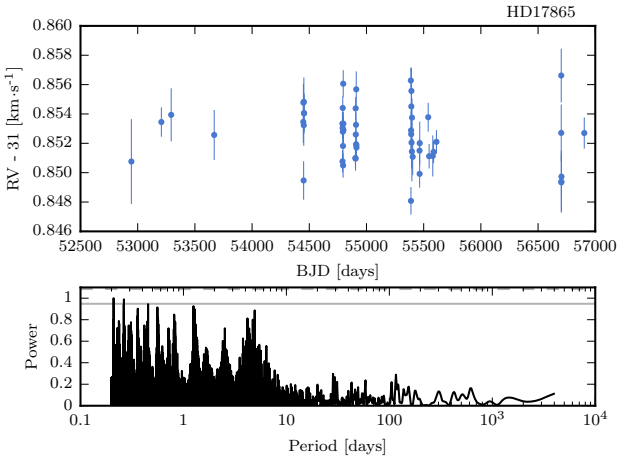

HD171587
RV + 24 [km·s⁻¹]
BJD [days]
Power
Period [days]

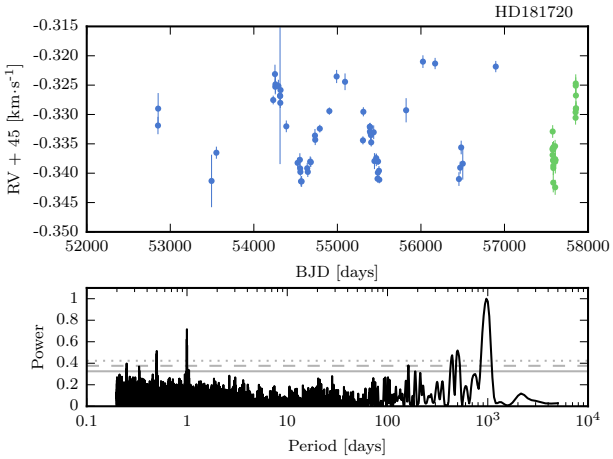

HD172568
RV + 45 [km·s⁻¹]
BJD [days]
Power
Period [days]

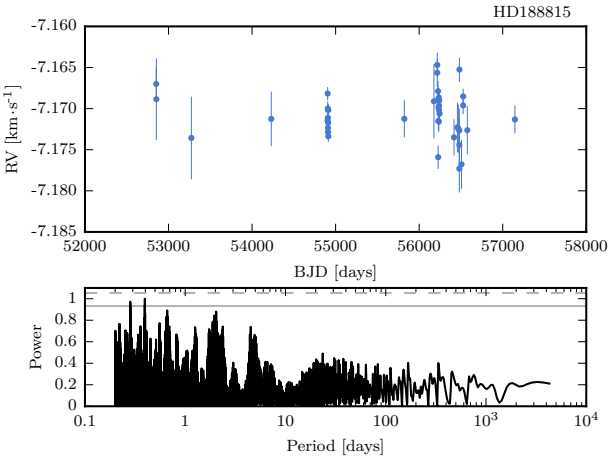

HD17548
RV + 15 [km·s⁻¹]
BJD [days]
Power
Period [days]

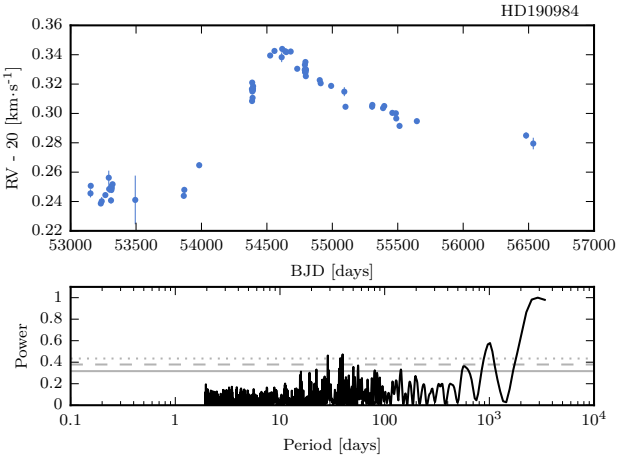

HD175607
RV + 91 [km·s⁻¹]
BJD [days]
Power
Period [days]

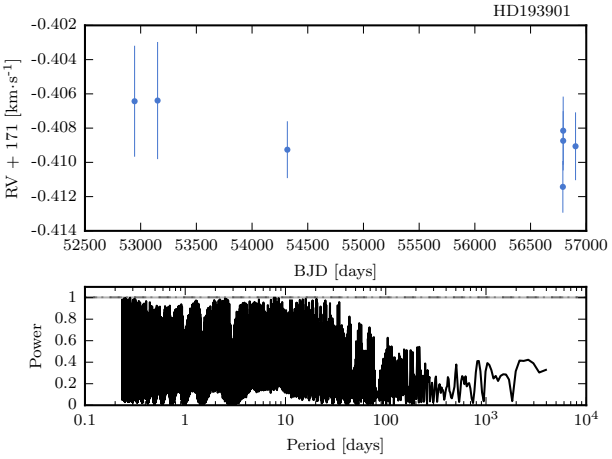

HD176666
RV + 31 [km·s⁻¹]
BJD [days]
Power
Period [days]

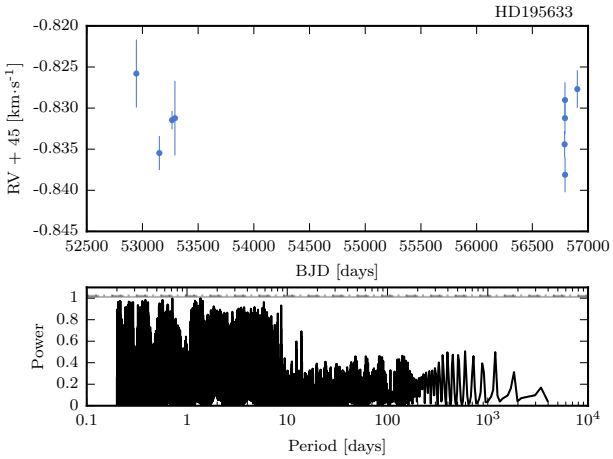

HD17865
RV - 31 [km·s⁻¹]
BJD [days]
Power
Period [days]

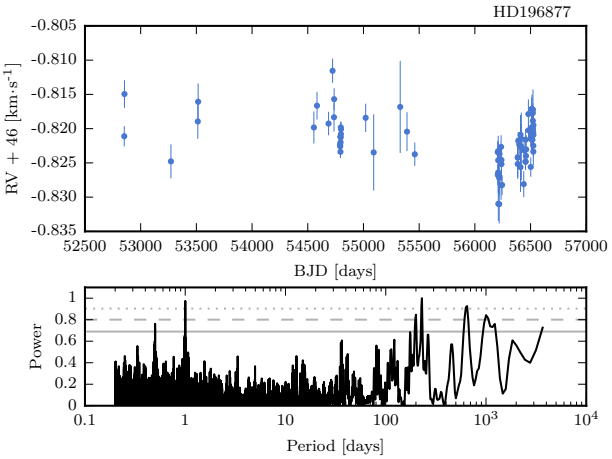

HD181720
RV + 45 [km·s⁻¹]
BJD [days]
Power
Period [days]

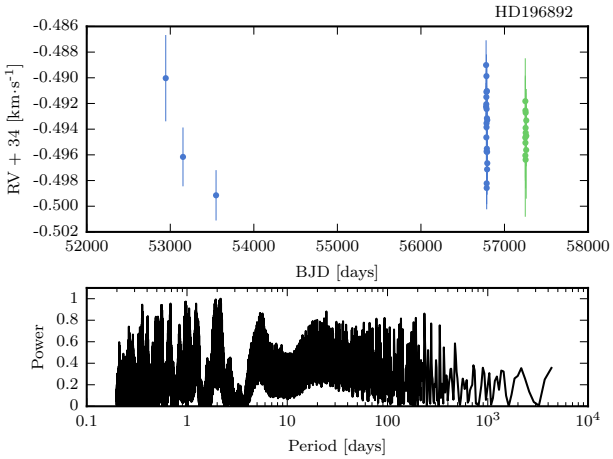

HD188815
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

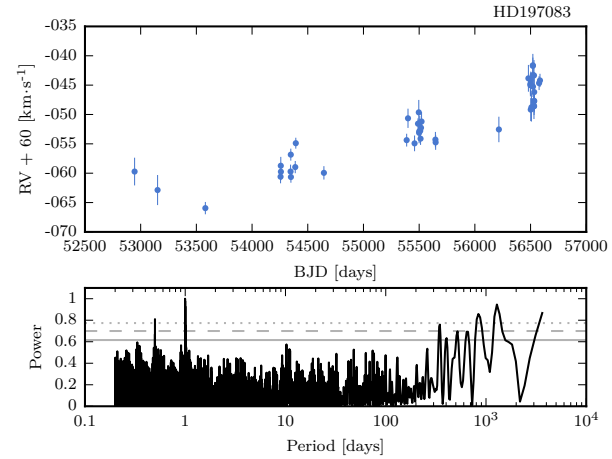

HD190984
RV - 20 [km·s⁻¹]
BJD [days]
Power
Period [days]

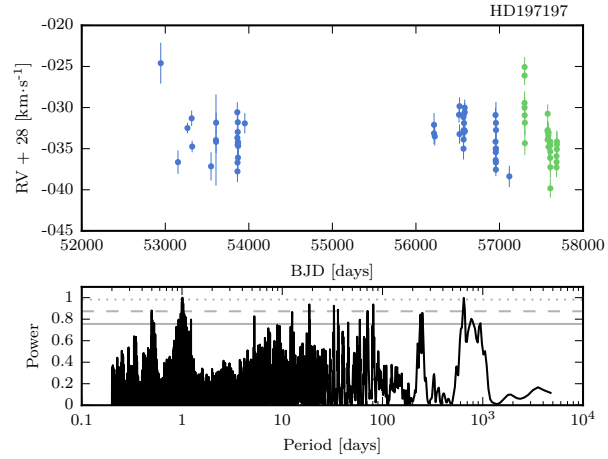

HD193901
RV + 171 [km·s⁻¹]
BJD [days]
Power
Period [days]

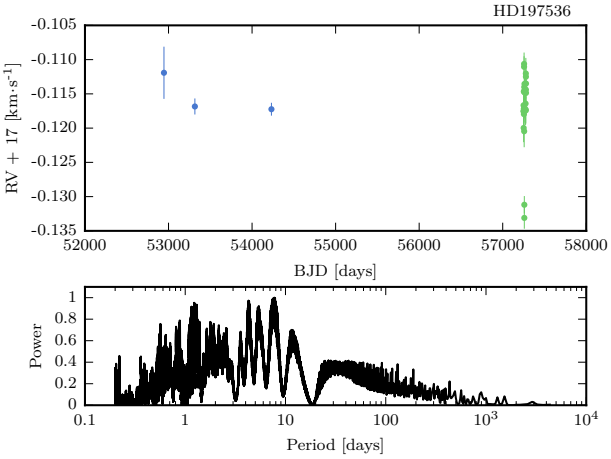

HD195633
RV + 45 [km·s⁻¹]
BJD [days]
Power
Period [days]

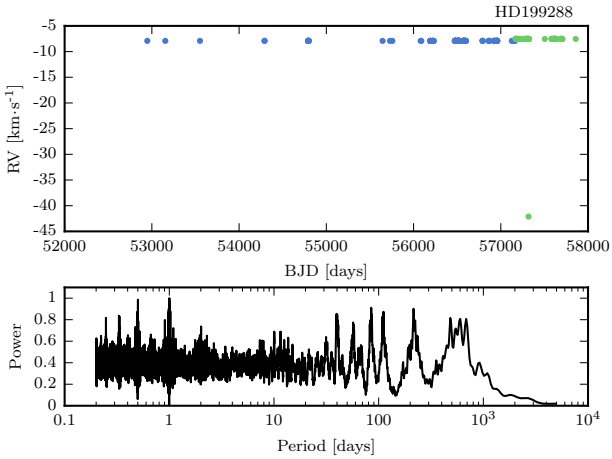

HD196877
RV + 46 [km·s⁻¹]
BJD [days]
Power
Period [days]

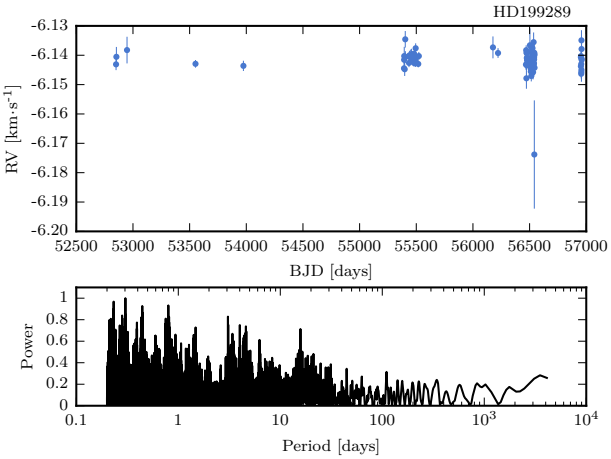

HD196892
RV + 34 [km·s⁻¹]
BJD [days]
Power
Period [days]

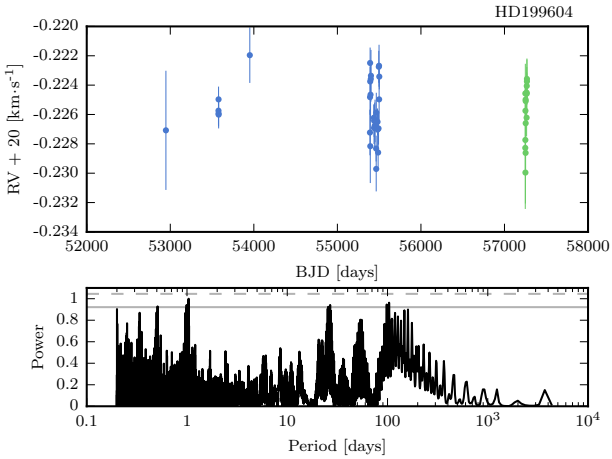

HD197083
RV + 60 [km·s⁻¹]
BJD [days]
Power
Period [days]

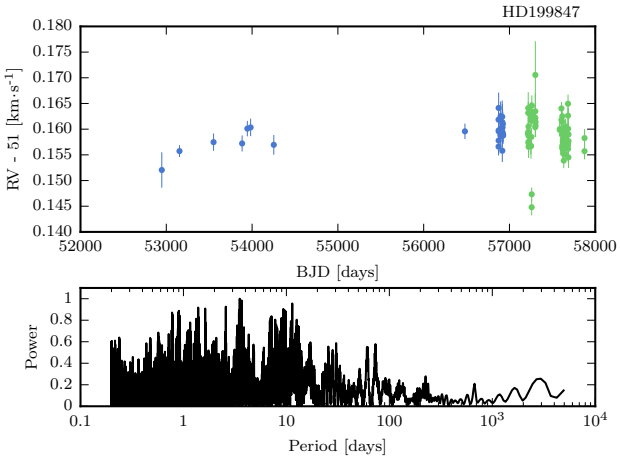

HD197197
RV + 28 [km·s⁻¹]
BJD [days]
Power
Period [days]

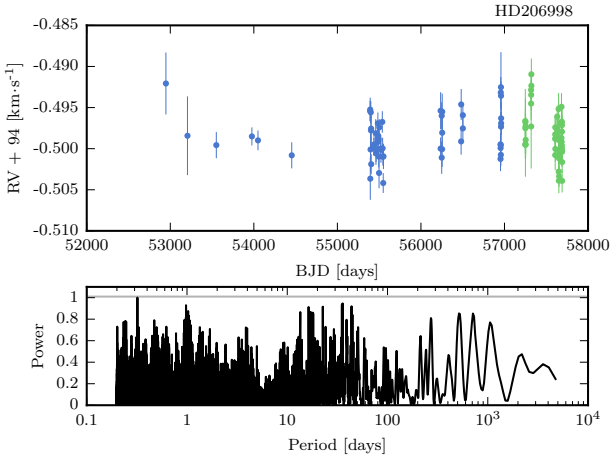

HD197536
RV + 17 [km·s⁻¹]
BJD [days]
Power
Period [days]

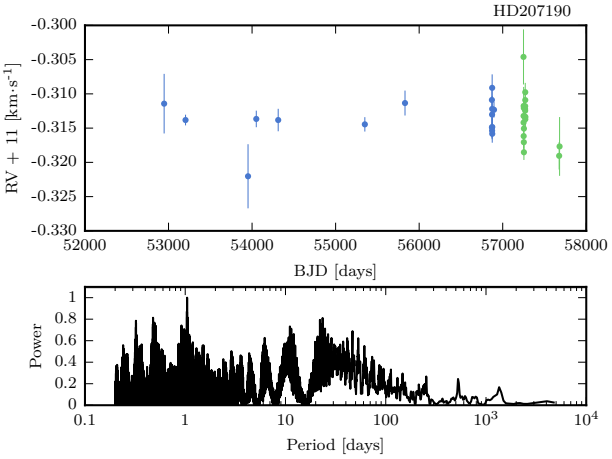

HD199288
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

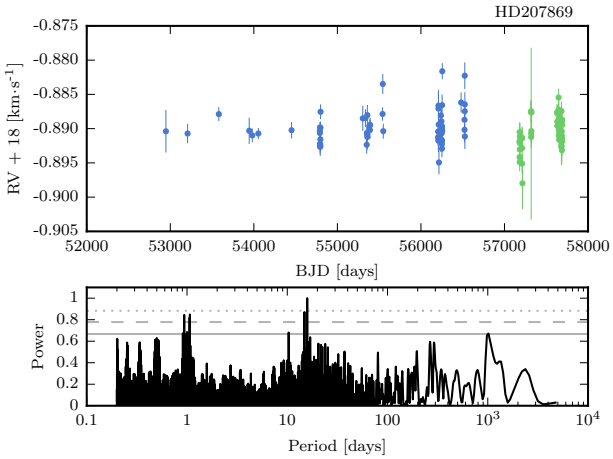

HD199289
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

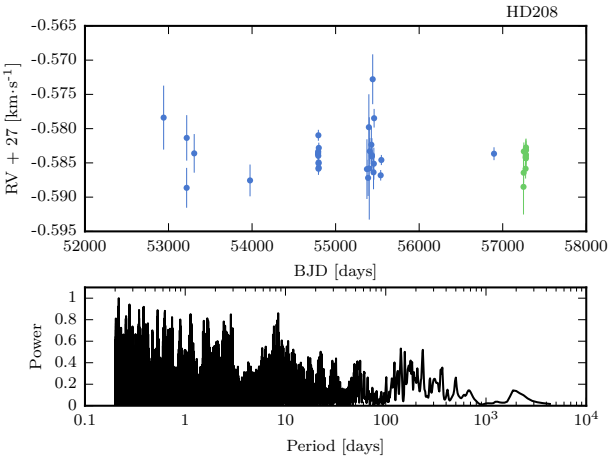

HD199604
RV + 20 [km·s⁻¹]
BJD [days]
Power
Period [days]

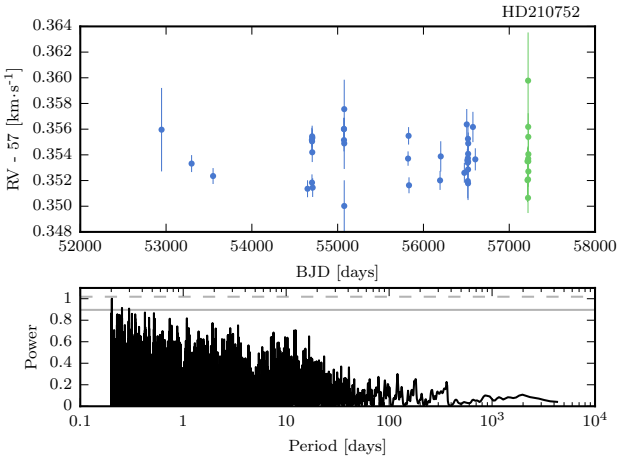

HD199847
RV - 51 [km·s⁻¹]
BJD [days]
Power
Period [days]

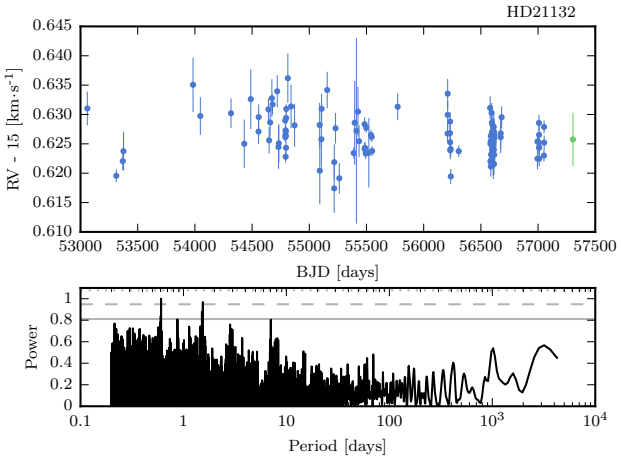

HD206998
RV + 94 [km·s⁻¹]
BJD [days]
Power
Period [days]

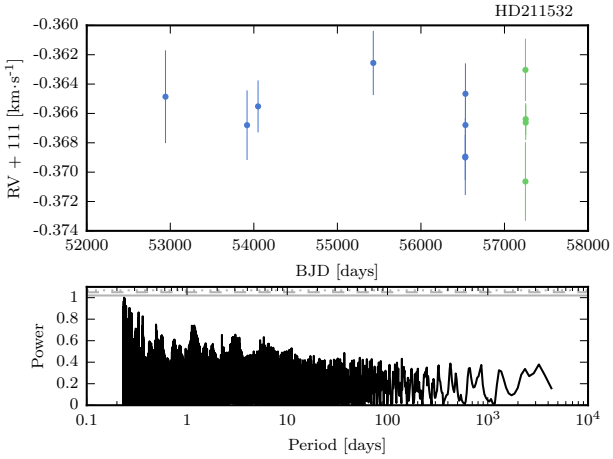

HD207190
RV + 11 [km·s⁻¹]
BJD [days]
Power
Period [days]

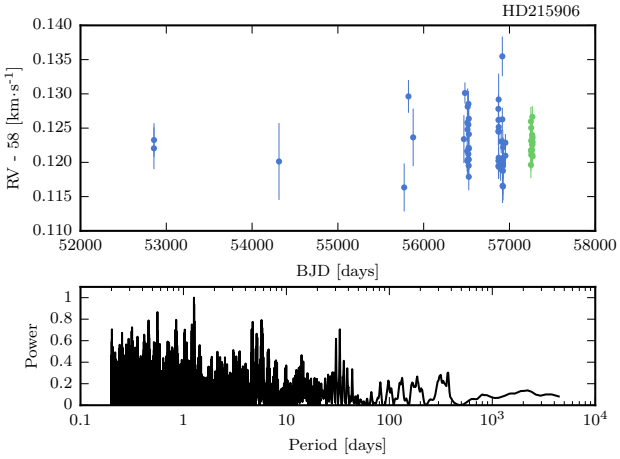

HD207869
RV + 18 [km·s⁻¹]
BJD [days]
Power
Period [days]

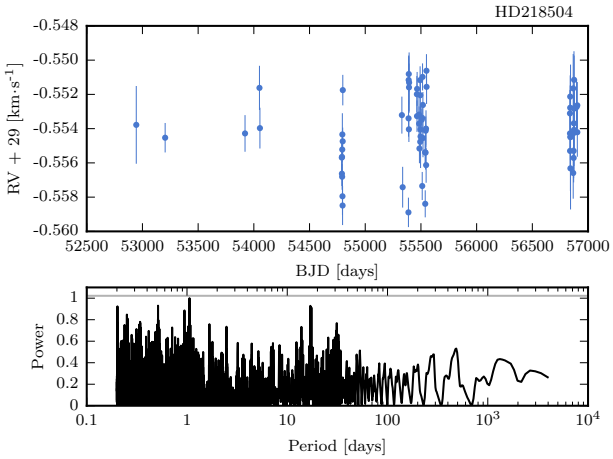

HD208
RV + 27 [km·s⁻¹]
BJD [days]
Power
Period [days]

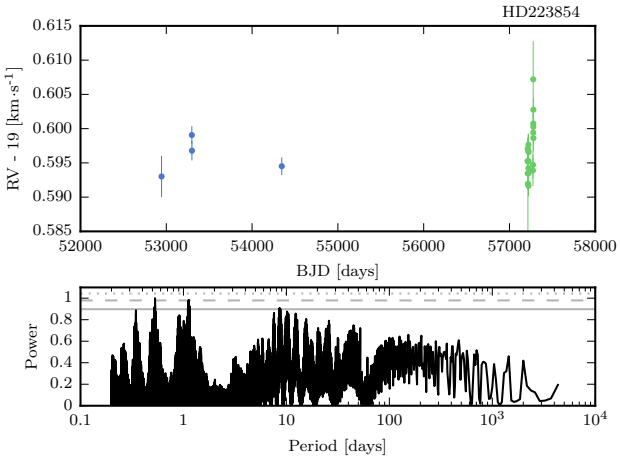

HD210752
RV - 57 [km·s⁻¹]
BJD [days]
Power
Period [days]

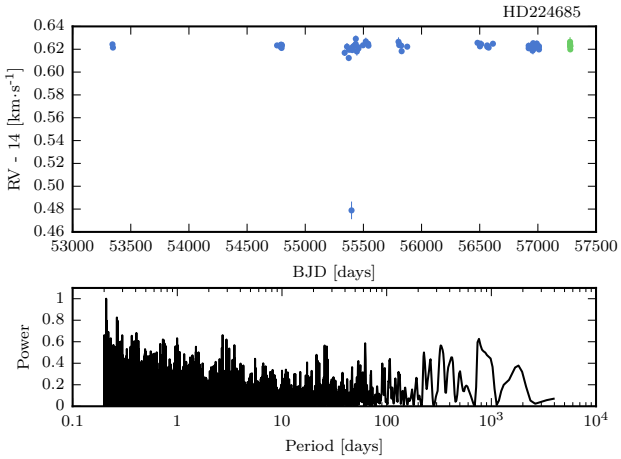

HD21132
RV - 15 [km·s⁻¹]
BJD [days]
Power
Period [days]

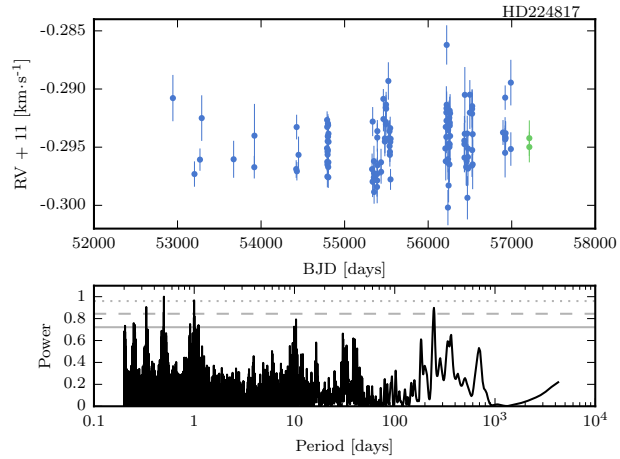

HD211532
RV + 111 [km·s⁻¹]
BJD [days]
Power
Period [days]

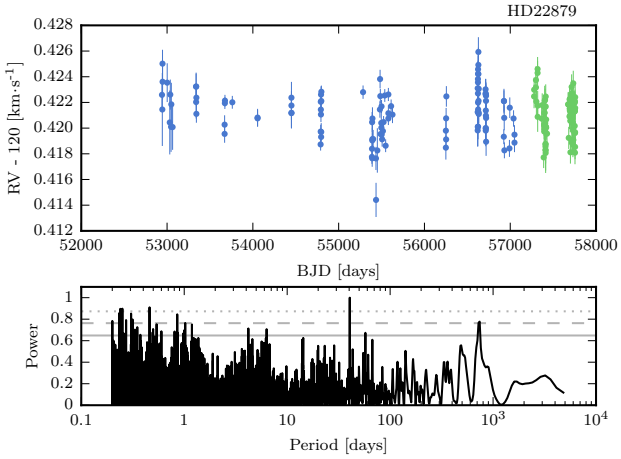

HD215906
RV - 58 [km·s⁻¹]
BJD [days]
Power
Period [days]

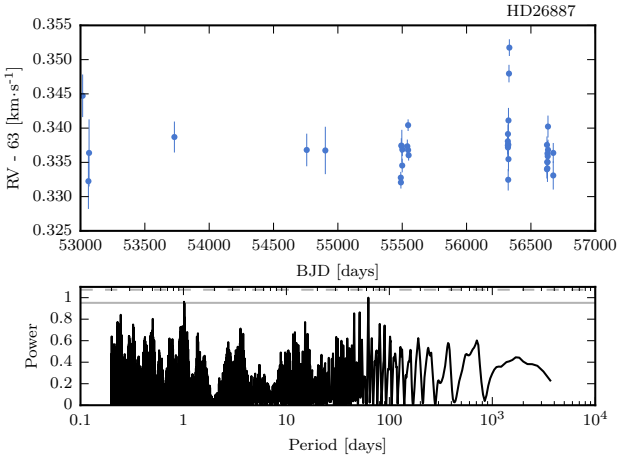

HD218504
RV + 29 [km·s⁻¹]
BJD [days]
Power
Period [days]

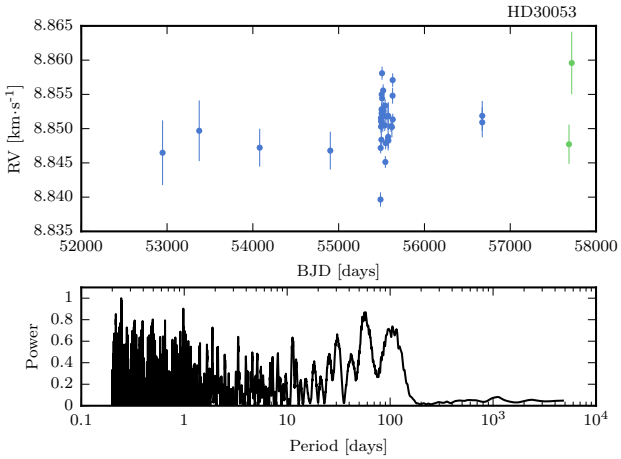

HD223854
RV - 19 [km·s⁻¹]
BJD [days]
Power
Period [days]

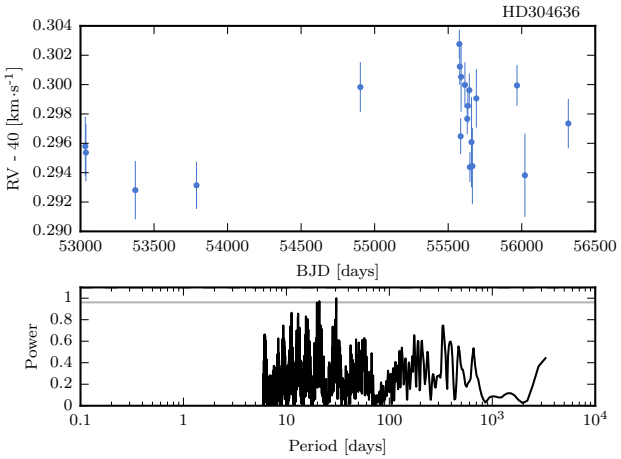

HD224685
RV - 14 [km·s⁻¹]
BJD [days]
Power
Period [days]

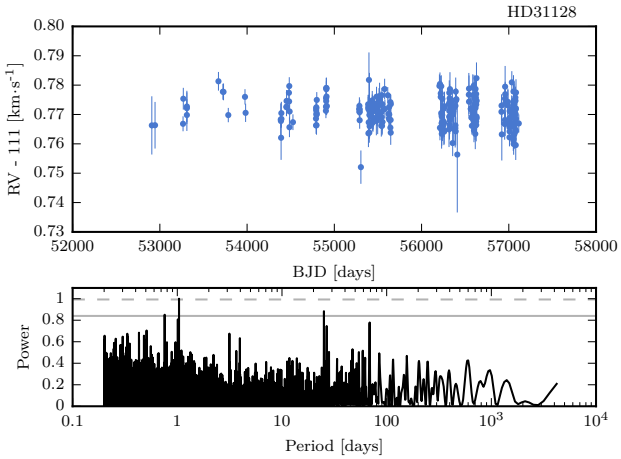

HD224817
RV + 11 [km·s⁻¹]
BJD [days]
Power
Period [days]

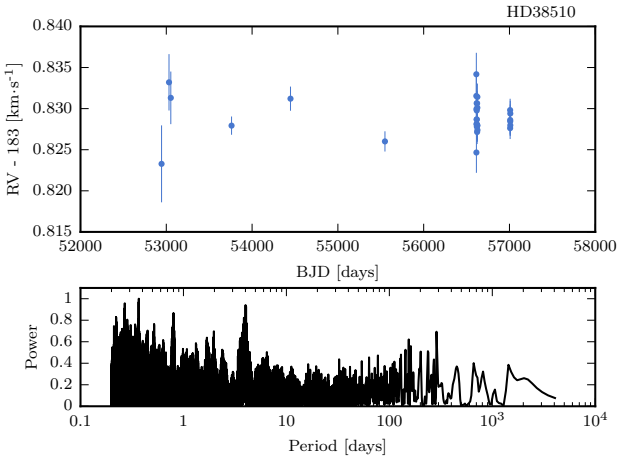

HD22879
RV - 120 [km·s^{-1}]
BJD [days]
Power
Period [days]

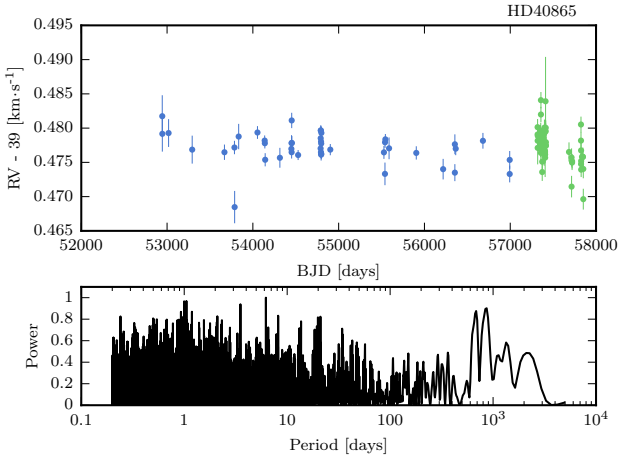

HD26887
RV - 63 [km·s^{-1}]
BJD [days]
Power
Period [days]

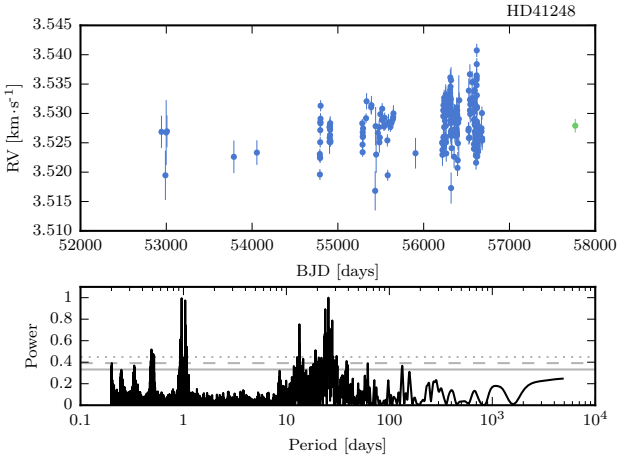

HD30053
RV [km·s^{-1}]
BJD [days]
Power
Period [days]

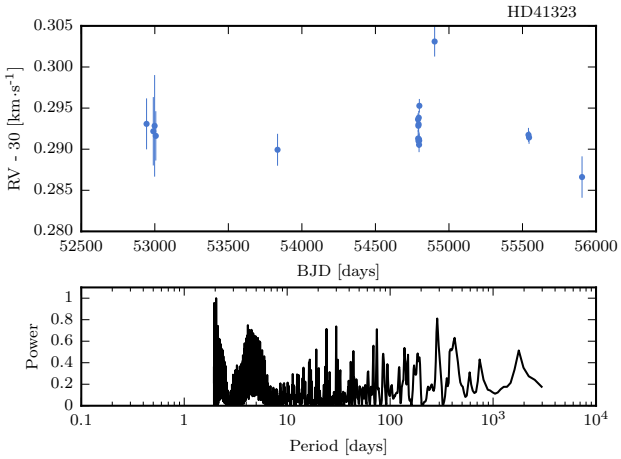

HD304636
RV - 40 [km·s^{-1}]
BJD [days]
Power
Period [days]

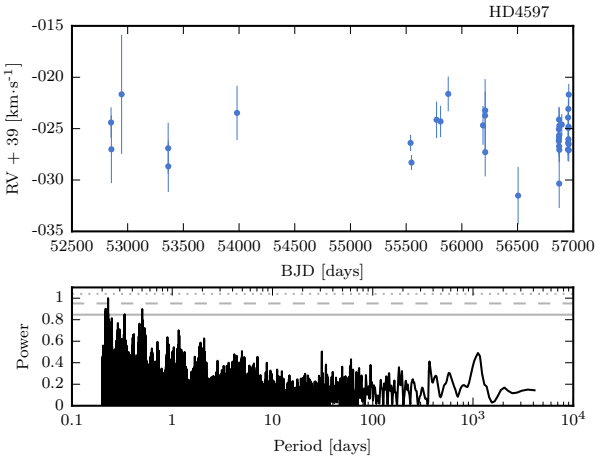

HD31128
RV - 111 [km·s^{-1}]
BJD [days]
Power
Period [days]

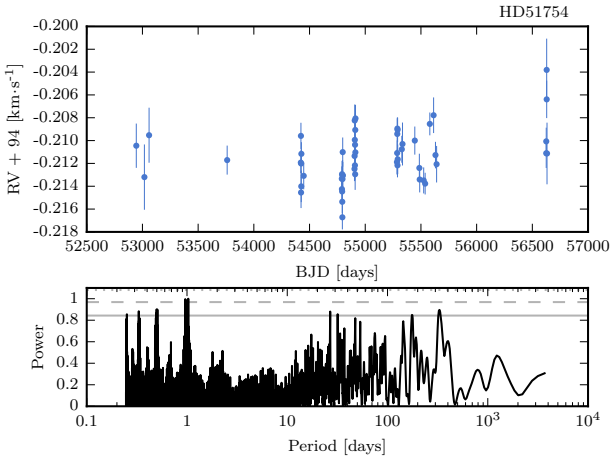

HD38510
RV - 183 [km·s^{-1}]
BJD [days]
Power
Period [days]

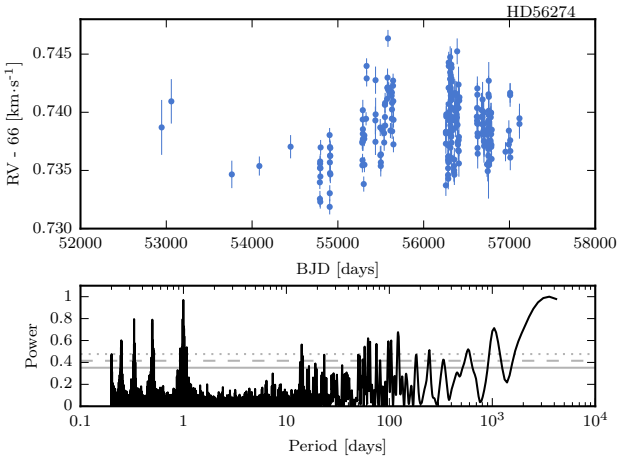

HD40865
RV - 39 [km·s⁻¹]
BJD [days]
Power
Period [days]

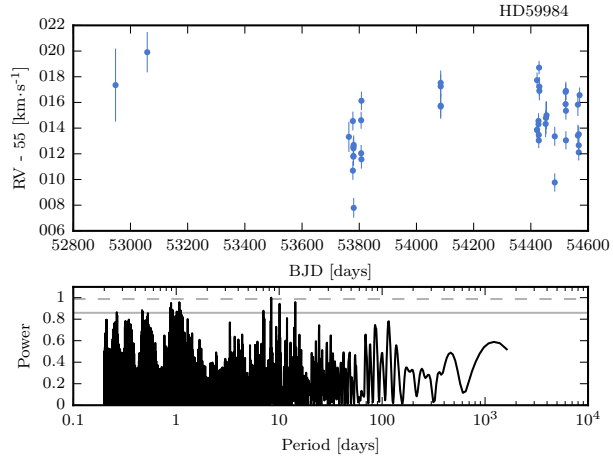

HD41248
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

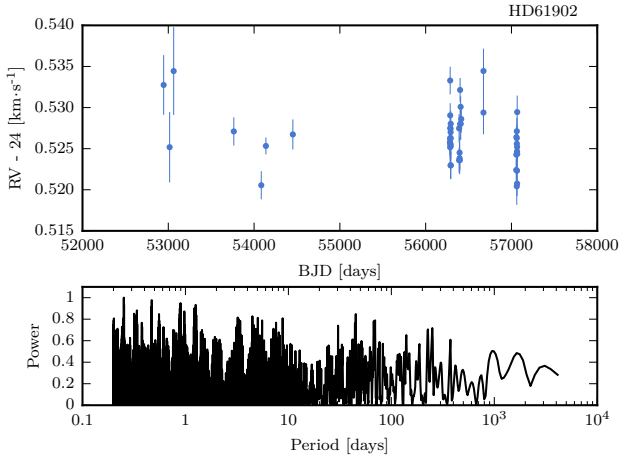

HD41323
RV - 30 [km·s⁻¹]
BJD [days]
Power
Period [days]

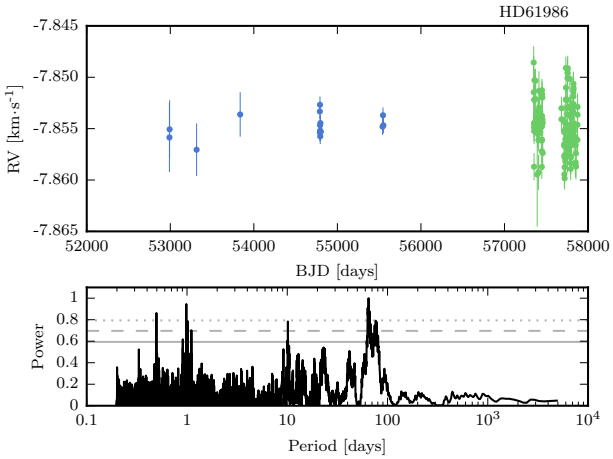

HD4597
RV + 39 [km·s⁻¹]
BJD [days]
Power
Period [days]

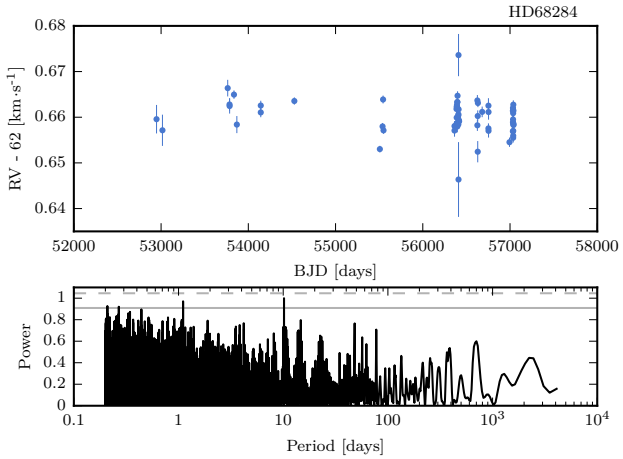

HD51754
RV + 94 [km·s⁻¹]
BJD [days]
Power
Period [days]

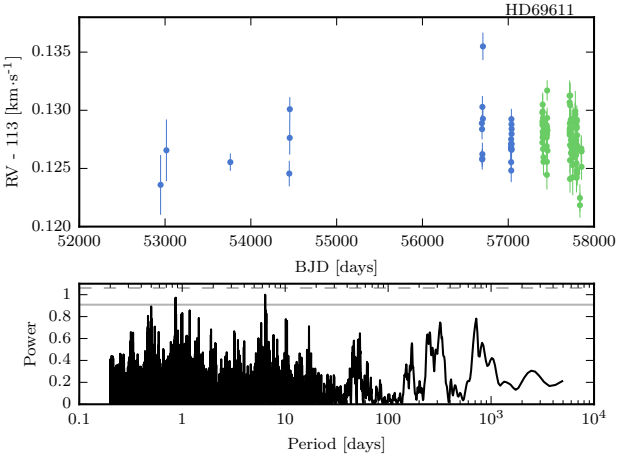

HD56274
RV - 66 [km·s⁻¹]
BJD [days]
Power
Period [days]

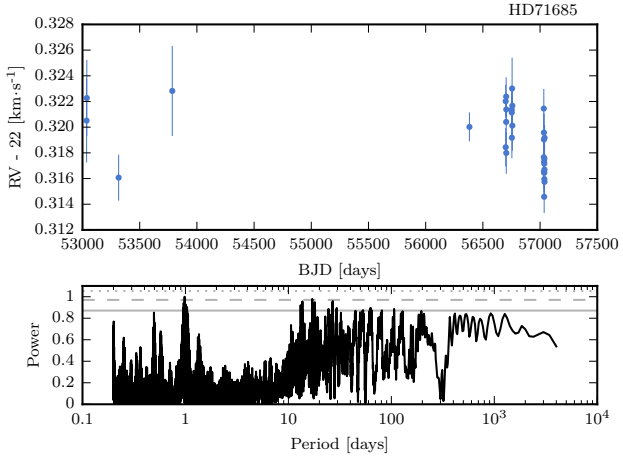

HD59984
RV - 55 [km·s⁻¹]
BJD [days]
Power
Period [days]

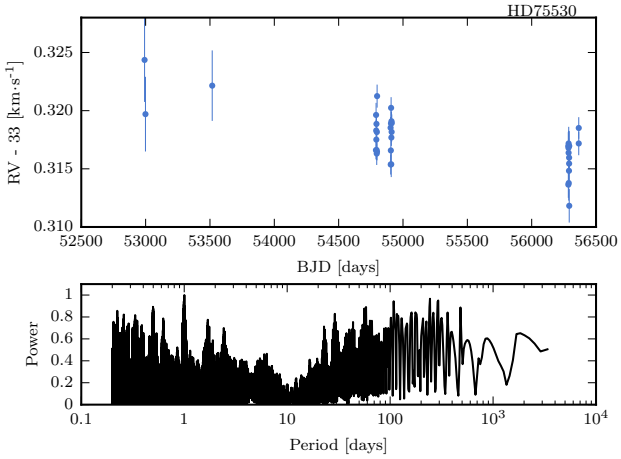

HD61902
RV - 24 [km·s⁻¹]
BJD [days]
Power
Period [days]

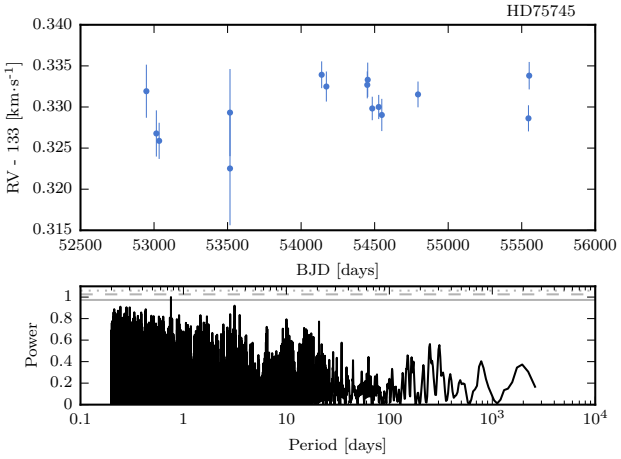

HD61986
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

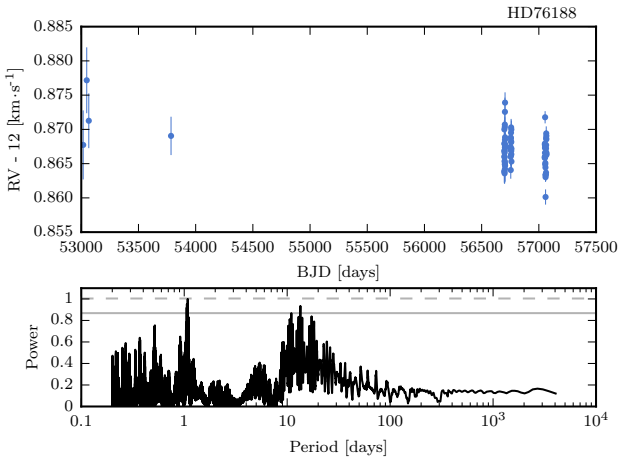

HD68284
RV - 62 [km·s⁻¹]
BJD [days]
Power
Period [days]

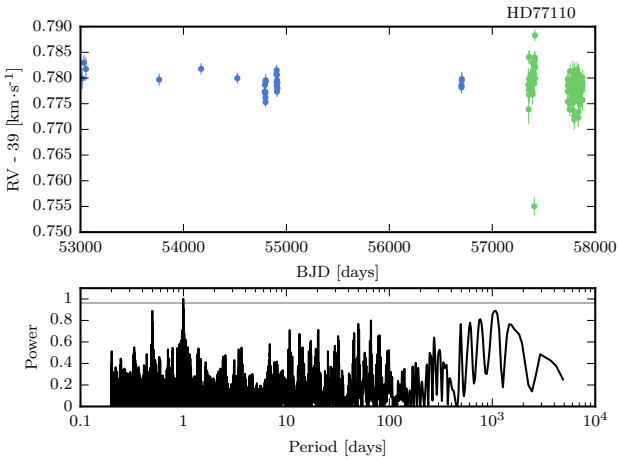

HD69611
RV - 113 [km·s⁻¹]
BJD [days]
Power
Period [days]

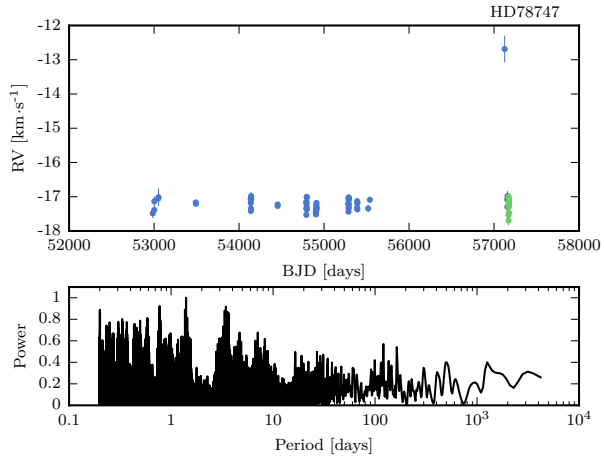

HD71685
RV - 22 [km·s⁻¹]
BJD [days]
Power
Period [days]

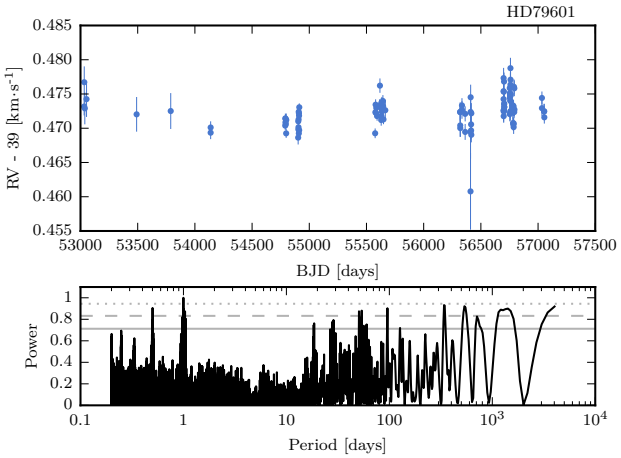

HD75530
RV - 33 [km·s⁻¹]
BJD [days]
Power
Period [days]

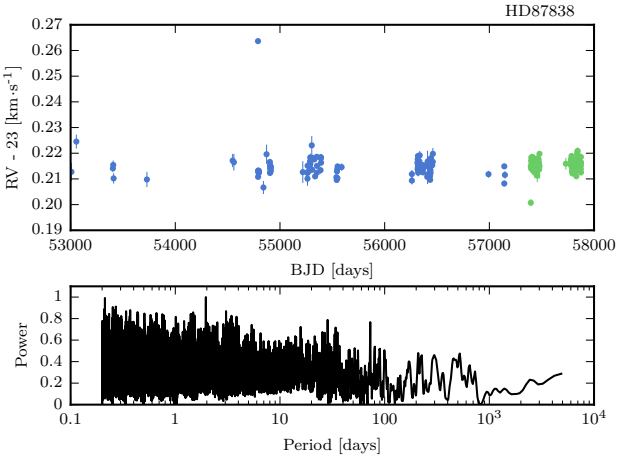

HD75745
RV - 133 [km·s⁻¹]
BJD [days]
Power
Period [days]

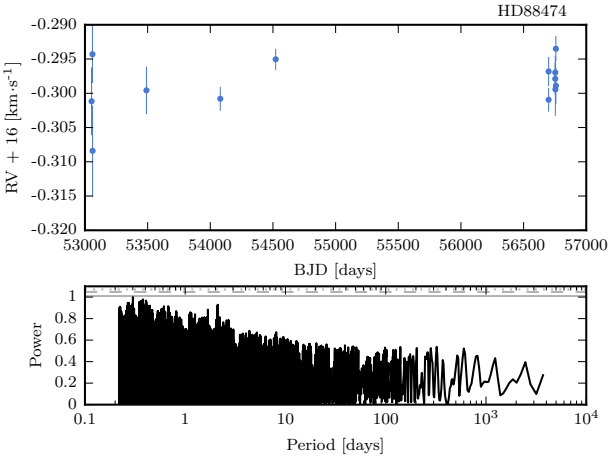

HD76188
RV - 12 [km·s⁻¹]
BJD [days]
Power
Period [days]

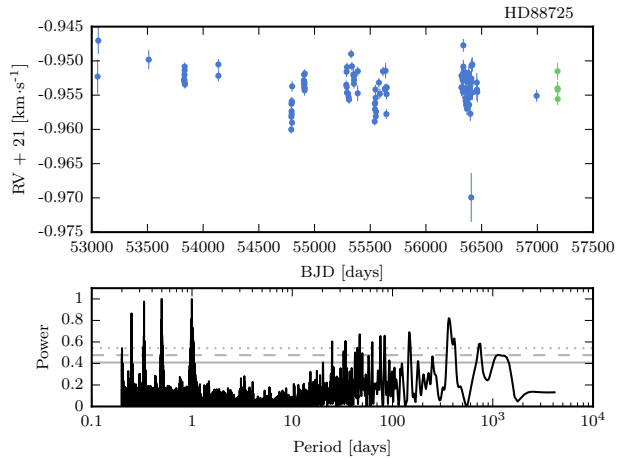

HD77110
RV - 39 [km·s⁻¹]
BJD [days]
Power
Period [days]

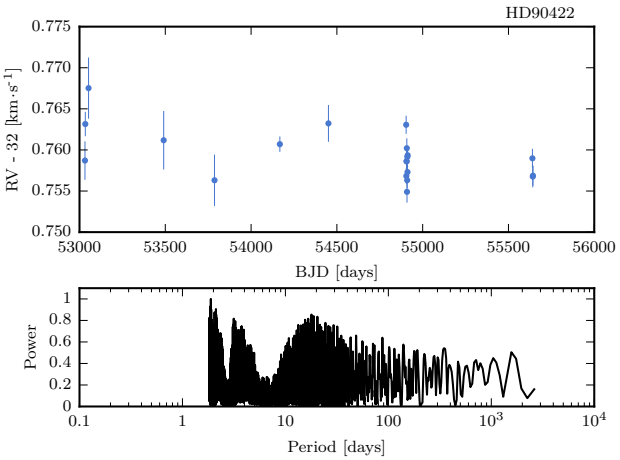

HD78747
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

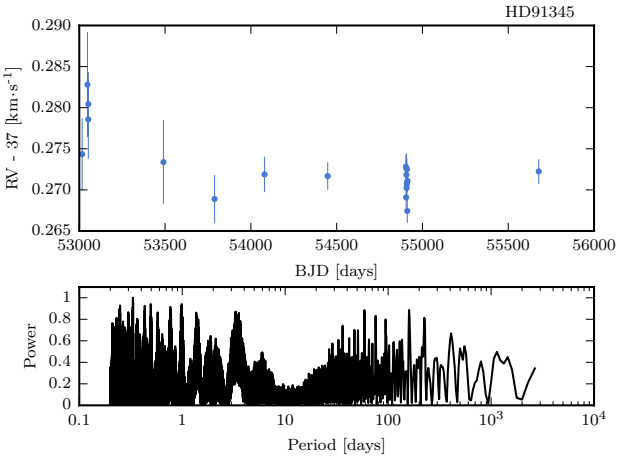

HD79601
RV - 39 [km·s⁻¹]
BJD [days]
Power
Period [days]

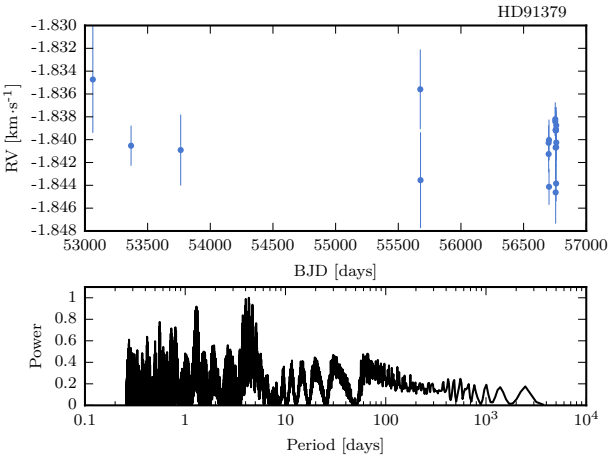

HD87838
RV - 23 [km·s⁻¹]
BJD [days]
Power
Period [days]

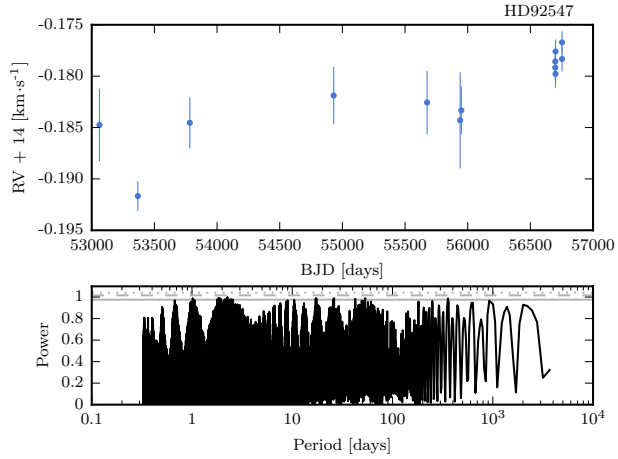

HD88474
RV + 16 [km·s⁻¹]
BJD [days]
Power
Period [days]

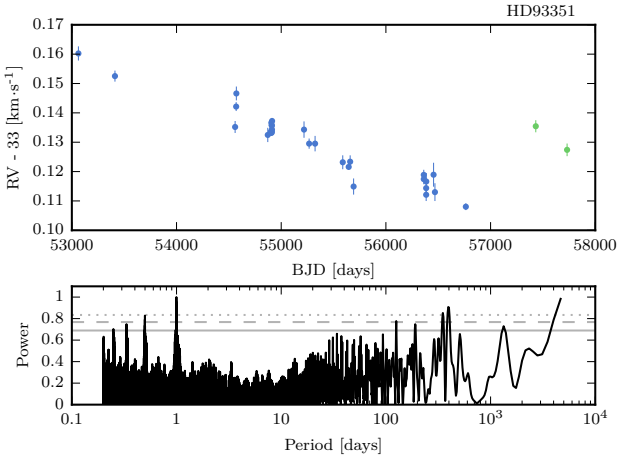

HD88725
RV + 21 [km·s⁻¹]
BJD [days]
Power
Period [days]

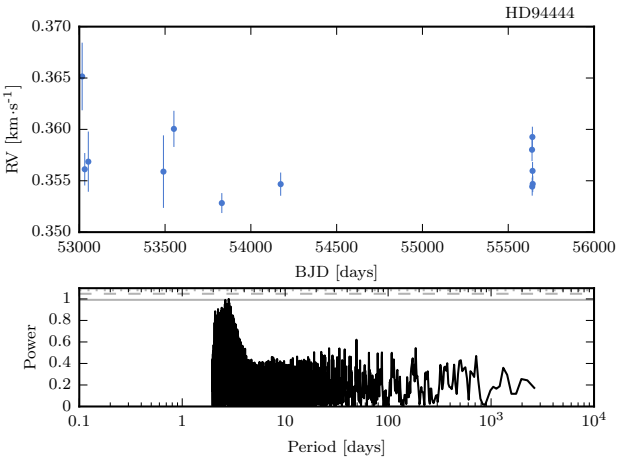

HD90422
RV - 32 [km·s⁻¹]
BJD [days]
Power
Period [days]

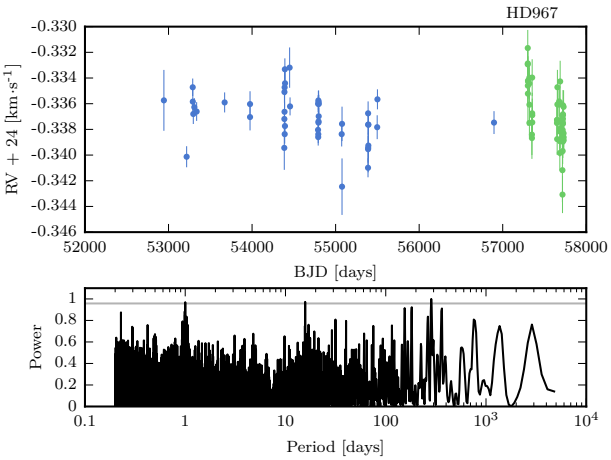

HD91345
RV - 37 [km·s⁻¹]
BJD [days]
Power
Period [days]

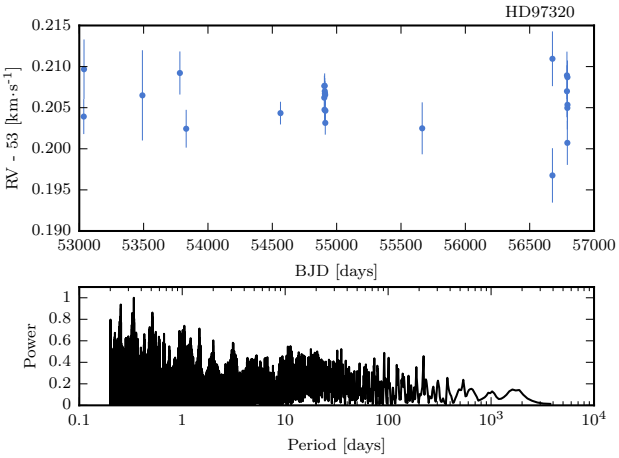

HD91379
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

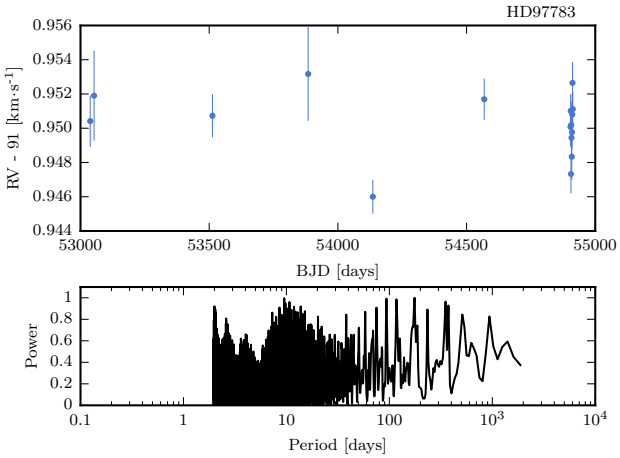

HD92547
RV + 14 [km·s⁻¹]
BJD [days]
Power
Period [days]

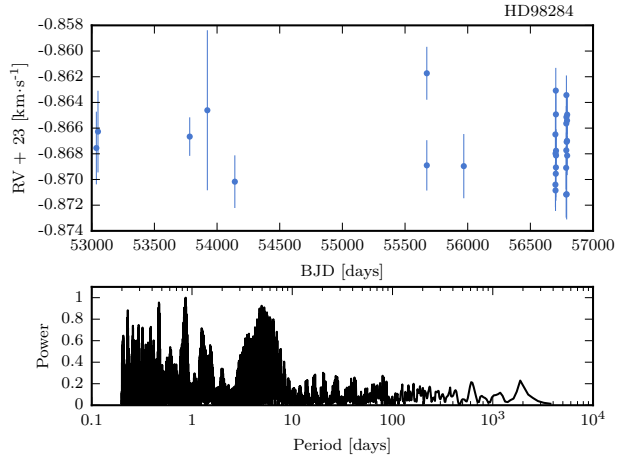

HD93351
RV - 33 [km·s⁻¹]
BJD [days]
Power
Period [days]

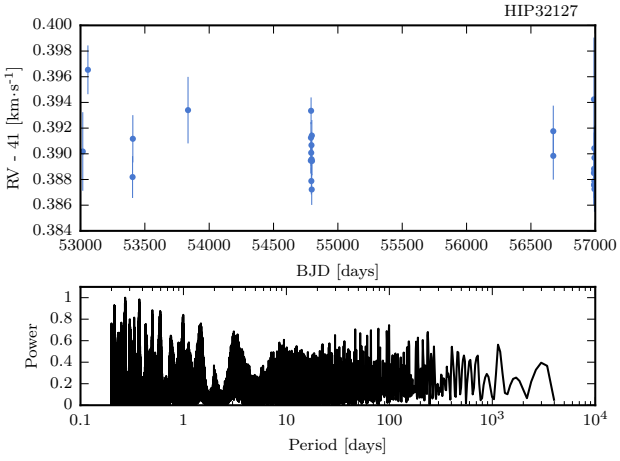

HD94444
RV [km·s⁻¹]
BJD [days]
Power
Period [days]

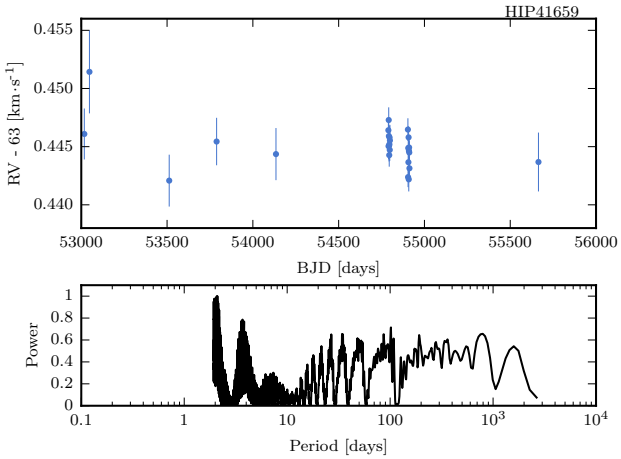

HD967
RV + 24 [km·s⁻¹]
BJD [days]
Power
Period [days]


HD97320
RV - 53 [km·s⁻¹]
BJD [days]
Power
Period [days]


HD97783
RV - 91 [km·s⁻¹]
BJD [days]
Power
Period [days]







In the pages below, Table C.1 presents the stellar parameters for the 109 targets in our sample. The stellar fundamental parameters were derived from the HARPS spectra, and the stellar masses using evolutionary models, as was described in Section 4.2. The Hipparcos astrometric parallaxes, re-derived by van Leeuwen (2007), were used to determine the absolute V-band magnitudes using the apparent visual magnitudes from Hipparcos. For some targets, the updated parallaxes included in the GAIA DR1 (Gaia Collaboration et al., 2016; Lindegren et al., 2016) were used instead.

The activity level of each star is represented by the weighted average of the log $R'_{\rm HK}$ timeseries, with the exceptions of HD 196877 and HD 304636 for which the chromospheric indicator was not derived. When the log $R'_{\rm HK}$ is available, the rotation period is estimated using the calibrations of Noyes et al. (1984b) and Mamajek and Hillenbrand (2008), considering the weighted average of the log $R'_{\rm HK}$ timeseries.

Table C.1: Observed and inferred stellar parameters for the stars in the metal-poor sample

Star	Mass $[M_\odot]$	$T_{\rm eff}$ [K]	$\log g$ [cgs]	[Fe/H]	$B-V$	V	π [mas]	M_V	$\log R'_{HK}$	P_{rot} [days]
BD-082534	0.74 ± 0.05	5405 ± 29	4.43 ± 0.04	-0.78 ± 0.02	0.66	9.51	17.06	5.67	-4.89	23.7 / 24.1
CD-571633	0.83 ± 0.06	5975 ± 41	4.46 ± 0.03	-0.85 ± 0.03	0.50	9.54	9.91	4.52	-4.87	7.7 / 7.7
HD 101612	1.02 ± 0.07	6281 ± 43	4.41 ± 0.03	-0.36 ± 0.03	0.47	7.54	18.78	3.91	-4.90	5.8 / 5.9
HD 101644	0.80 ± 0.06	5678 ± 18	4.58 ± 0.02	-0.56 ± 0.01	0.68	9.25	15.84	5.25	-4.79	22.6 / 22.2
HD 104800	0.78 ± 0.06	5697 ± 25	4.47 ± 0.02	-0.79 ± 0.02	0.58	9.22	14.24	4.99	-4.91	15.3 / 15.7
HD 107094	0.79 ± 0.06	5564 ± 17	4.54 ± 0.02	-0.51 ± 0.01	0.66	9.13	18.83	5.50	-4.82	21.1 / 20.8
HD 108564	0.65 ± 0.05	4818 ± 69	4.67 ± 0.17	-0.97 ± 0.07	0.98	9.45	36.78	7.28	-4.86	41.5 / 41.5
HD 109684	0.96 ± 0.07	5992 ± 18	4.38 ± 0.02	-0.34 ± 0.01	0.55	8.73	12.90	4.28	-4.97	13.2 / 13.8
HD 111515	0.76 ± 0.05	5398 ± 18	4.47 ± 0.02	-0.61 ± 0.01	0.76	8.10	30.71	5.54	-4.93	34.9 / 35.9
HD 111777	0.80 ± 0.06	5666 ± 19	4.46 ± 0.03	-0.68 ± 0.01	0.61	8.46	20.67	5.04	-4.91	18.7 / 19.1
HD 11397	0.79 ± 0.06	5565 ± 26	4.50 ± 0.04	-0.55 ± 0.02	0.61	9.02	19.03	5.42	-4.89	18.1 / 18.4
HD 114076	0.75 ± 0.06	5069 ± 52	4.32 ± 0.09	-0.47 ± 0.04	0.82	9.39	19.72	5.86	-4.98	41.4 / 43.4
HD 119173	0.88 ± 0.06	5779 ± 44	4.26 ± 0.04	-0.62 ± 0.03	0.56	8.83	18.43	5.16	-4.87	12.7 / 12.8
HD 119949	1.01 ± 0.07	6359 ± 36	4.47 ± 0.04	-0.41 ± 0.02	0.45	8.16	11.96	3.55	-4.96	5.0 / 5.2
HD 121004	0.79 ± 0.06	5687 ± 26	4.48 ± 0.03	-0.71 ± 0.02	0.60	9.04	16.70	5.15	-4.90	17.4 / 17.6
HD 123651	0.87 ± 0.06	5926 ± 30	4.55 ± 0.03	-0.48 ± 0.02	0.53	8.19	18.93	4.58	-4.85	9.8 / 9.8
HD 126681	0.70 ± 0.05	5561 ± 32	4.71 ± 0.03	-1.14 ± 0.02	0.61	9.32	21.04	5.94	-4.84	17.0 / 16.9
HD 126793	0.84 ± 0.06	5910 ± 31	4.46 ± 0.03	-0.71 ± 0.02	0.52	8.24	18.53	4.58	-4.92	9.8 / 10.0
HD 126803	0.77 ± 0.05	5477 ± 23	4.50 ± 0.04	-0.61 ± 0.02	0.68	8.94	19.18	5.35	-4.92	26.6 / 27.2
HD 128340	0.92 ± 0.07	6259 ± 40	4.64 ± 0.02	-0.55 ± 0.03	0.47	8.88	11.62	4.21	-4.89	5.8 / 5.9

Table C.1: continued.

Star	Mass $[M_\odot]$	$T_{\rm eff}$ [K]	$\log g$ [cgs]	[Fe/H]	$B-V$	V	π [mas]	M_V	$\log R'_{HK}$	P_{rot} [days]
HD 128571	0.99 ± 0.07	6159 ± 47	4.40 ± 0.02	-0.37 ± 0.03	0.50	7.82	17.32	4.01	-4.93	8.2 / 8.5
HD 129229	1.15 ± 0.08	5872 ± 21	3.89 ± 0.04	-0.42 ± 0.02	0.57	8.41	7.10	2.67	-5.10	17.1 / 18.8
HD 131653	0.73 ± 0.05	5324 ± 26	4.54 ± 0.04	-0.66 ± 0.02	0.72	9.50	20.17	6.02	-4.96	32.3 / 33.7
HD 133633	0.81 ± 0.06	5571 ± 19	4.48 ± 0.04	-0.45 ± 0.01	0.65	8.79	17.51	5.01	-4.92	23.5 / 24.1
HD 134088	0.79 ± 0.06	5675 ± 22	4.46 ± 0.03	-0.75 ± 0.02	0.59	8.00	26.62	5.13	-4.90	16.2 / 16.5
HD 134440	0.64 ± 0.04	4987 ± 48	4.80 ± 0.08	-1.32 ± 0.03	0.79	9.43	35.14	7.16	-4.90	36.1 / 36.6
HD 137676	0.88 ± 0.06	5253 ± 18	3.93 ± 0.03	-0.53 ± 0.01	0.76	7.69	17.30	3.88	-5.08	40.2 / 43.8
HD 141624	0.91 ± 0.06	5871 ± 30	4.40 ± 0.03	-0.38 ± 0.02	0.56	8.19	20.16	4.71	-4.93	13.6 / 14.0
HD 144589	1.22 ± 0.08	6372 ± 37	4.28 ± 0.03	-0.05 ± 0.03	0.54	9.80	4.59	3.11	-5.04	13.0 / 13.9
HD 145344	0.91 ± 0.07	6143 ± 41	4.39 ± 0.04	-0.68 ± 0.03	0.48	8.42	11.84	3.79	-4.92	6.7 / 6.9
HD 145417	0.64 ± 0.04	5006 ± 53	4.82 ± 0.12	-1.23 ± 0.04	0.82	7.54	72.01	6.83	-4.87	37.3 / 37.5
HD 147518	0.81 ± 0.06	5626 ± 30	4.40 ± 0.03	-0.63 ± 0.02	0.62	9.38	13.25	4.99	-4.89	19.4 / 19.6
HD 148211	0.89 ± 0.06	5922 ± 21	4.34 ± 0.03	-0.62 ± 0.01	0.53	7.69	19.32	4.12	-4.95	10.9 / 11.3
HD 148816	0.85 ± 0.06	5908 ± 25	4.39 ± 0.02	-0.71 ± 0.02	0.53	7.28	23.41	4.13	-4.94	10.8 / 11.1
HD 149747	1.11 ± 0.08	5823 ± 35	3.95 ± 0.04	-0.34 ± 0.03	0.63	9.17	4.92	2.63	-5.01	23.0 / 24.5
HD 150177	1.03 ± 0.07	6216 ± 28	4.18 ± 0.03	-0.58 ± 0.02	0.44	6.34	24.96	3.33	-4.96	4.3 / 4.5
HD 167300	0.91 ± 0.06	5837 ± 20	4.30 ± 0.03	-0.45 ± 0.01	0.50	9.23	8.97	3.99	-5.00	8.8 / 9.2
HD 171028	1.08 ± 0.08	5671 ± 16	3.84 ± 0.03	-0.48 ± 0.01	0.64	8.29	8.86	3.03	-5.05	25.1 / 27.1
HD 171587	0.75 ± 0.05	5412 ± 15	4.59 ± 0.02	-0.64 ± 0.01	0.67	8.48	24.15	5.39	-4.75	20.2 / 19.7
HD 172568	0.85 ± 0.06	5728 ± 22	4.58 ± 0.03	-0.37 ± 0.02	0.63	8.54	20.88	5.14	-4.90	20.7 / 21.1
HD 17548	0.90 ± 0.06	6011 ± 26	4.44 ± 0.02	-0.53 ± 0.02	0.48	8.18	18.19	4.48	-4.91	6.6 / 6.7

Table C.1: continued.

Star	Mass $[M_\odot]$	$T_{\rm eff}$ [K]	$\log g$ [cgs]	[Fe/H]	$B-V$	V	π [mas]	M_V	$\log R'_{HK}$	P_{rot} [days]
HD 175607	0.75 ± 0.05	5392 ± 17	4.51 ± 0.03	-0.61 ± 0.01	0.70	8.61	22.09	5.33	-4.92	28.9 / 29.7
HD 176666	0.93 ± 0.07	6103 ± 34	4.63 ± 0.03	-0.37 ± 0.03	0.52	8.29	17.02	4.44	-4.89	9.5 / 9.6
HD 17865	0.89 ± 0.06	5877 ± 24	4.32 ± 0.03	-0.57 ± 0.02	0.55	8.19	15.60	4.16	-4.96	13.1 / 13.6
HD 181720	0.90 ± 0.06	5794 ± 17	4.25 ± 0.02	-0.53 ± 0.01	0.58	7.86	17.22	4.04	-4.99	16.7 / 17.7
HD 188815	0.98 ± 0.07	6217 ± 49	4.34 ± 0.03	-0.53 ± 0.03	0.42	7.50	19.11	3.91	-4.90	3.3 / 3.3
HD 190984	1.08 ± 0.08	6007 ± 25	4.02 ± 0.03	-0.49 ± 0.02	0.54	8.73	6.84	2.91	-5.04	13.0 / 13.9
HD 193901	0.74 ± 0.05	5611 ± 34	4.41 ± 0.05	-1.07 ± 0.03	0.55	8.67	22.78	5.46	-4.90	12.2 / 12.4
HD 195633	1.00 ± 0.07	6154 ± 37	4.25 ± 0.05	-0.51 ± 0.03	0.51	8.52	10.07	3.54	-4.99	9.6 / 10.1
HD 196877	0.70 ± 0.20	4479 ± 184	4.04 ± 0.53	-0.72 ± 0.05	1.33	8.82	82.31	8.40	—	— / —
HD 196892	0.83 ± 0.06	6072 ± 56	4.50 ± 0.03	-0.89 ± 0.03	0.50	8.23	16.15	4.27	-4.90	7.9 / 8.0
HD 197083	0.84 ± 0.06	5735 ± 16	4.50 ± 0.02	-0.45 ± 0.01	0.62	9.19	13.75	4.88	-4.92	19.8 / 20.3
HD 197197	0.94 ± 0.07	5812 ± 16	4.20 ± 0.02	-0.46 ± 0.01	0.62	8.07	14.50	3.88	-5.03	22.1 / 23.6
HD 197536	0.97 ± 0.07	6105 ± 24	4.39 ± 0.03	-0.41 ± 0.02	0.52	8.20	14.15	3.95	-4.97	10.3 / 10.8
HD 199288	0.82 ± 0.06	5746 ± 22	4.46 ± 0.03	-0.63 ± 0.02	0.59	6.52	45.17	4.79	-4.90	16.3 / 16.6
HD 199289	0.78 ± 0.06	5925 ± 42	4.62 ± 0.03	-0.98 ± 0.03	0.52	8.30	18.95	4.69	-4.89	9.5 / 9.6
HD 199604	0.86 ± 0.06	5817 ± 22	4.34 ± 0.03	-0.62 ± 0.02	0.53	8.61	14.82	4.46	-4.93	10.8 / 11.1
HD 199847	0.90 ± 0.06	5763 ± 20	4.22 ± 0.02	-0.54 ± 0.01	0.59	8.81	12.73	4.33	-5.01	18.1 / 19.3
HD 206998	0.88 ± 0.06	5822 ± 26	4.24 ± 0.03	-0.69 ± 0.02	0.56	8.67	11.32	3.94	-4.97	14.1 / 14.8
HD 207190	1.01 ± 0.07	6182 ± 27	4.33 ± 0.03	-0.42 ± 0.02	0.50	7.66	16.79	3.79	-4.96	8.5 / 8.9
HD 207869	0.80 ± 0.06	5527 ± 21	4.50 ± 0.05	-0.45 ± 0.02	0.64	8.97	21.55	5.64	-4.93	22.6 / 23.2
HD 208	0.92 ± 0.07	5914 ± 20	4.47 ± 0.03	-0.31 ± 0.02	0.54	8.19	18.25	4.50	-4.92	11.5 / 11.8

Table C.1: continued.

Star	Mass $[M_\odot]$	$T_{\rm eff}$ [K]	$\log g$ [cgs]	[Fe/H]	$B-V$	V	π [mas]	M_V	$\log R'_{HK}$	P_{rot} [days]
HD 210752	0.86 ± 0.06	5951 ± 21	4.53 ± 0.03	-0.58 ± 0.02	0.52	7.40	27.64	4.61	-4.87	9.2 / 9.3
HD 21132	0.97 ± 0.07	6243 ± 34	4.60 ± 0.05	-0.37 ± 0.02	0.49	7.86	18.17	4.16	-4.90	7.2 / 7.3
HD 211532	0.74 ± 0.05	5199 ± 22	4.42 ± 0.04	-0.54 ± 0.02	0.79	9.27	19.81	5.75	-5.00	40.0 / 42.3
HD 215906	1.00 ± 0.07	6259 ± 58	4.56 ± 0.07	-0.28 ± 0.04	0.47	7.76	17.58	3.99	-4.92	6.0 / 6.2
HD 218504	0.91 ± 0.06	5962 ± 29	4.34 ± 0.03	-0.55 ± 0.02	0.54	8.11	15.27	4.03	-4.96	12.0 / 12.5
HD 223854	1.05 ± 0.07	6080 ± 30	4.08 ± 0.03	-0.54 ± 0.02	0.46	8.07	11.98	3.46	-5.00	5.8 / 6.2
HD 224685	0.81 ± 0.06	5504 ± 30	4.47 ± 0.06	-0.40 ± 0.02	0.64	9.22	18.84	5.60	-4.90	21.8 / 22.1
HD 224817	0.89 ± 0.06	5894 ± 22	4.36 ± 0.02	-0.53 ± 0.02	0.55	8.40	13.68	4.08	-4.96	13.0 / 13.5
HD 22879	0.81 ± 0.06	5884 ± 33	4.52 ± 0.03	-0.81 ± 0.02	0.56	6.67	39.12	4.63	-4.91	12.8 / 13.1
HD 26887	0.94 ± 0.07	6016 ± 28	4.46 ± 0.02	-0.35 ± 0.02	0.52	8.46	17.97	4.73	-4.88	9.3 / 9.3
HD 30053	1.00 ± 0.07	6139 ± 41	4.51 ± 0.03	-0.22 ± 0.03	0.54	8.14	17.69	4.38	-4.88	11.0 / 11.1
HD 304636	—	3636	—	-0.15	1.49	9.46	95.58	9.37	—	— / —
HD 31128	0.74 ± 0.05	6096 ± 67	4.90 ± 0.05	-1.39 ± 0.04	0.41	9.14	15.00	5.02	-4.88	2.8 / 2.9
HD 38510	0.85 ± 0.06	5914 ± 37	4.32 ± 0.03	-0.81 ± 0.02	0.50	8.25	15.49	4.20	-4.94	8.2 / 8.5
HD 40865	0.86 ± 0.06	5722 ± 17	4.49 ± 0.02	-0.38 ± 0.01	0.62	8.61	19.64	5.08	-4.91	19.8 / 20.2
HD 41248	0.86 ± 0.06	5713 ± 21	4.49 ± 0.03	-0.37 ± 0.01	0.61	8.81	19.09	5.21	-4.89	18.2 / 18.4
HD 41323	0.87 ± 0.06	5756 ± 22	4.56 ± 0.03	-0.31 ± 0.02	0.64	8.71	19.50	5.16	-4.85	20.5 / 20.5
HD 4597	0.94 ± 0.07	6025 ± 31	4.43 ± 0.03	-0.39 ± 0.02	0.54	7.85	20.51	4.41	-4.92	11.5 / 11.8
HD 51754	0.85 ± 0.06	5848 ± 24	4.49 ± 0.02	-0.55 ± 0.02	0.56	9.01	12.56	4.50	-4.92	13.3 / 13.6
HD 56274	0.83 ± 0.06	5734 ± 22	4.51 ± 0.03	-0.54 ± 0.02	0.57	7.79	30.95	5.24	-4.85	13.3 / 13.3
HD 59984	0.93 ± 0.07	5962 ± 27	4.18 ± 0.02	-0.69 ± 0.02	0.48	5.93	35.82	3.70	-4.97	7.1 / 7.4

Table C.1: continued.

Star	Mass $[M_\odot]$	$T_{\rm eff}$ [K]	$\log g$ [cgs]	[Fe/H]	$B-V$	V	π [mas]	M_V	$\log R'_{HK}$	P_{rot} [days]
HD 61902	0.94 ± 0.07	6209 ± 30	4.38 ± 0.03	-0.62 ± 0.02	0.48	8.25	12.71	3.77	-4.92	6.7 / 6.8
HD 61986	0.87 ± 0.06	5725 ± 20	4.48 ± 0.04	-0.34 ± 0.02	0.64	8.66	19.16	5.07	-4.93	22.4 / 22.9
HD 68284	1.02 ± 0.07	5933 ± 26	4.08 ± 0.03	-0.50 ± 0.02	0.53	7.75	13.14	3.34	-5.05	12.1 / 13.0
HD 69611	0.87 ± 0.06	5762 ± 25	4.31 ± 0.03	-0.58 ± 0.02	0.58	7.74	20.50	4.30	-4.97	16.3 / 17.1
HD 71685	0.92 ± 0.07	6038 ± 30	4.58 ± 0.04	-0.37 ± 0.02	0.53	8.35	17.91	4.62	-4.87	10.1 / 10.1
HD 75530	0.75 ± 0.05	5311 ± 29	4.48 ± 0.07	-0.54 ± 0.02	0.74	9.19	17.86	5.45	-4.96	34.3 / 35.7
HD 75745	0.86 ± 0.06	5885 ± 35	4.29 ± 0.03	-0.78 ± 0.03	0.57	9.44	7.55	3.83	-4.93	14.7 / 15.1
HD 76188	1.05 ± 0.08	5989 ± 46	4.08 ± 0.04	-0.44 ± 0.03	0.52	7.16	18.03	3.44	-5.02	10.7 / 11.4
HD 77110	0.83 ± 0.06	5717 ± 20	4.48 ± 0.02	-0.50 ± 0.01	0.58	8.85	16.28	4.91	-4.92	15.6 / 16.0
HD 78747	0.83 ± 0.06	5788 ± 20	4.44 ± 0.02	-0.67 ± 0.01	0.51	7.74	24.53	4.69	-4.92	8.9 / 9.1
HD 79601	0.86 ± 0.06	5834 ± 25	4.37 ± 0.04	-0.60 ± 0.02	0.53	8.02	17.70	4.26	-4.95	11.0 / 11.4
HD 87838	0.95 ± 0.07	6118 ± 33	4.47 ± 0.03	-0.40 ± 0.02	0.52	7.72	21.25	4.36	-4.92	9.7 / 9.9
HD 88474	1.21 ± 0.08	6122 ± 40	3.91 ± 0.03	-0.48 ± 0.03	0.54	8.47	6.51	2.54	-5.13	14.1 / 15.6
HD 88725	0.80 ± 0.06	5654 ± 17	4.49 ± 0.03	-0.64 ± 0.01	0.60	7.73	28.24	4.98	-4.90	17.3 / 17.6
HD 90422	1.00 ± 0.07	6085 ± 33	4.14 ± 0.03	-0.62 ± 0.02	0.52	8.24	10.15	3.27	-4.99	10.5 / 11.1
HD 91345	0.74 ± 0.05	5658 ± 39	4.52 ± 0.04	-1.04 ± 0.03	0.54	9.03	16.71	5.14	-4.87	10.9 / 11.0
HD 91379	1.01 ± 0.07	6164 ± 46	4.41 ± 0.04	-0.29 ± 0.03	0.50	8.16	18.85	4.54	-4.91	8.0 / 8.1
HD 92547	0.94 ± 0.07	6020 ± 33	4.45 ± 0.03	-0.37 ± 0.02	0.54	8.11	17.93	4.38	-4.93	11.7 / 12.0
HD 93351	0.83 ± 0.06	5408 ± 25	4.41 ± 0.05	-0.23 ± 0.02	0.71	9.14	17.47	5.35	-4.98	31.7 / 33.2
HD 94444	0.90 ± 0.06	5998 ± 27	4.34 ± 0.03	-0.61 ± 0.02	0.52	8.10	17.38	4.30	-4.94	9.9 / 10.2
HD 967	0.77 ± 0.05	5568 ± 17	4.53 ± 0.02	-0.68 ± 0.01	0.61	8.38	23.50	5.24	-4.91	18.6 / 19.0

Table C.1: continued.

Star	Mass $[M_\odot]$	$T_{\rm eff}$ [K]	$\log g$ [cgs]	[Fe/H]	$B-V$	V	π [mas]	M_V	$\log R'_{HK}$	P_{rot} [days]
HD 97320	0.81 ± 0.06	6162 ± 52	4.57 ± 0.04	-1.05 ± 0.03	0.50	8.15	18.36	4.47	-4.86	7.5 / 7.5
HD 97783	0.79 ± 0.06	5682 ± 24	4.50 ± 0.02	-0.73 ± 0.02	0.63	9.04	15.02	4.92	-4.90	20.6 / 21.0
HD 98284	0.81 ± 0.06	5913 ± 39	4.52 ± 0.04	-0.84 ± 0.03	0.54	8.30	20.04	4.81	-4.89	11.1 / 11.2
HIP32127	0.74 ± 0.05	5302 ± 28	4.44 ± 0.06	-0.64 ± 0.02	0.69	9.47	17.26	5.66	-4.97	29.3 / 30.6
HIP41659	0.75 ± 0.06	5197 ± 29	4.38 ± 0.09	-0.53 ± 0.02	0.77	9.51	19.44	5.95	-4.98	37.9 / 39.8

Table C.2, in the next pages, provides some descriptive statistics of all the timeseries for the stars in the metal-poor sample. For each star, the table lists the total number of RV observations, the number of nightly-binned observations, the average RV uncertainty $\bar{\sigma}_i$, the weighted standard deviation of all the RV measurements, and finally the timespan of observations for each of the targets.

Table C.2: Descriptive statistics of the RV timeseries for the 109 stars in the metal-poor sample.

Star	N	N_{binned}	$\bar{\sigma}_i$ [m·s^{-1}]	s_{RV} [m·s^{-1}]	Δt [days]	Star	N	N_{binned}	$\bar{\sigma}_i$ [m·s^{-1}]	s_{RV} [m·s^{-1}]	Δt [days]
BD-082534	16	16	1.93	1.81	1920	HD 199289	83	67	1.92	1.97	4108
CD-571633	7	7	2.79	1.96	1502	HD 199604	45	37	1.42	6.52	4321
HD 101612	44	22	1.54	1.14	3774	HD 199847	105	59	1.49	6.18	4927
HD 101644	23	23	2.06	2.65	1888	HD 206998	107	75	1.57	7.67	4742
HD 104800	64	62	1.59	2.16	3984	HD 207190	35	23	1.75	10.08	4733
HD 107094	88	65	1.62	767.84	4465	HD 207869	113	77	1.24	6.66	4742
HD 108564	32	24	1.52	1.57	3361	HD 208	41	37	2.05	6.59	4337
HD 109684	25	25	1.52	2.26	2337	HD 210752	47	36	1.00	4.61	4275
HD 111515	75	48	1.00	6.87	4811	HD 21132	117	92	2.20	2.60	4245
HD 111777	186	105	1.08	5.98	4807	HD 211532	12	8	1.87	5.77	4311
HD 11397	70	68	1.40	6.54	4790	HD 215906	67	42	1.91	8.88	4413
HD 114076	137	94	1.37	1.91	4052	HD 218504	70	61	1.10	2.04	3959
HD 119173	128	98	1.59	1.62	2959	HD 223854	24	15	1.98	10.20	4334
HD 119949	106	81	1.50	2.23	3614	HD 224685	66	54	1.75	4.38	3936
HD 121004	27	18	1.70	6.49	4169	HD 224817	123	101	1.15	2.39	4266
HD 123651	48	30	1.65	5.07	4076	HD 22879	202	138	0.84	6.99	4815
HD 126681	20	19	2.37	2.86	3309	HD 26887	39	28	1.63	3.86	3657
HD 126793	106	87	1.27	1.95	3913	HD 30053	38	35	1.50	3.74	4768
HD 126803	91	55	1.32	5.01	4819	HD 304636	20	20	1.73	2.87	3284
HD 128340	5	5	2.68	1.61	1031	HD 31128	264	200	3.59	3.73	4206
HD 128571	33	21	2.35	7.97	4147	HD 38510	31	19	1.64	1.37	4064
HD 129229	13	13	1.31	1.66	1422	HD 40865	100	73	1.21	7.02	4907
HD 131653	91	52	1.36	3.25	4059	HD 41248	229	161	1.26	3.45	4822
HD 133633	76	45	1.13	6.44	5004	HD 41323	19	19	1.73	2.03	2960
HD 134088	128	80	0.98	2.81	4032	HD 4597	37	28	1.74	1.46	4108
HD 134440	50	42	2.61	2.41	3392	HD 51754	54	52	1.32	2.11	3686
HD 137676	106	85	0.84	6.27	4758	HD 56274	235	165	0.80	2.69	4170
HD 141624	59	34	1.24	7.10	4359	HD 59984	45	23	0.76	2.05	1622
HD 144589	40	38	4.77	56.42	3744	HD 61902	48	32	1.68	2.54	4120
HD 145344	11	11	2.38	1.41	1761	HD 61986	155	89	1.11	5.91	4884
HD 145417	83	59	1.11	1.49	3744	HD 68284	63	44	1.20	2.78	4091
HD 147518	40	31	1.48	5.12	4062	HD 69611	122	68	0.97	7.59	4906
HD 148211	76	71	1.35	2.67	4337	HD 71685	32	20	1.50	2.08	4003
HD 148816	28	27	0.99	1.68	2312	HD 75530	37	31	1.16	1.64	3373
HD 149747	9	7	3.30	3.07	1327	HD 75745	15	14	2.09	2.32	2602
HD 150177	64	42	1.17	3.00	3322	HD 76188	74	32	1.19	1.70	4049
HD 167300	38	26	1.50	2.10	3697	HD 77110	152	90	1.13	6.55	4859
HD 171028	72	63	1.24	38.08	4565	HD 78747	57	46	0.79	1.30	4173

Table C.2: continued.

Star	N	N_{binned}	$\bar{\sigma}_i$ [m·s^{-1}]	s_{RV} [m·s^{-1}]	Δt [days]	Star	N	N_{binned}	$\bar{\sigma}_i$ [m·s^{-1}]	s_{RV} [m·s^{-1}]	Δt [days]
HD 171587	51	51	1.11	3.16	3585	HD 79601	108	84	1.00	1.75	4022
HD 172568	57	40	1.31	9.27	5025	HD 87838	209	143	1.25	8.44	4867
HD 17548	44	39	1.36	2.98	4360	HD 88474	13	10	3.10	2.46	3704
HD 175607	129	116	1.08	2.61	3912	HD 88725	111	92	0.79	2.78	4129
HD 176666	38	25	2.54	7.35	4395	HD 90422	20	20	1.61	2.29	2610
HD 17865	54	52	1.14	1.73	3959	HD 91345	19	18	2.34	2.05	2660
HD 181720	80	68	1.25	7.84	5003	HD 91379	21	13	2.07	1.67	3697
HD 188815	42	36	1.75	2.10	4293	HD 92547	13	10	2.29	4.08	3690
HD 190984	59	59	2.06	32.55	3383	HD 93351	33	31	1.79	9.54	4668
HD 193901	7	6	2.19	1.46	3958	HD 94444	12	12	1.65	2.24	2625
HD 195633	9	8	2.42	2.93	3958	HD 967	93	66	1.00	6.44	4784
HD 196877	80	64	1.69	2.58	3673	HD 97320	25	21	2.03	1.74	3758
HD 196892	35	21	1.97	7.60	4315	HD 97783	16	16	1.35	1.76	1874
HD 197083	45	38	1.41	6.00	3638	HD 98284	30	18	1.85	1.67	3756
HD 197197	78	55	1.23	8.03	4742	HIP32127	28	22	1.47	1.91	3977
HD 197536	28	16	1.40	7.64	4333	HIP41659	26	26	1.37	1.38	2646
HD 199288	184	121	0.71	6.43	4911						

References

Adibekyan, V. Zh.; Delgado Mena, E.; Sousa, S. G.; Santos, N. C.; Israelian, G.; González Hernández, J. I.; Mayor, M., and Hakobyan, A. A. (2012a) Exploring the α-Enhancement of Metal-Poor Planet-Hosting Stars. The Kepler and HARPS Samples in: *Astronomy and Astrophysics*, **547**: A36 DOI: 10.1051/0004-6361/201220167 (see pp. 27, 118)

Adibekyan, V. Zh.; Figueira, P.; Santos, N. C.; Mortier, A.; Mordasini, C.; Delgado Mena, E.; Sousa, S. G.; Correia, A. C. M.; Israelian, G., and Oshagh, M. (2013) Orbital and Physical Properties of Planets and Their Hosts: New Insights on Planet Formation and Evolution in: *Astronomy and Astrophysics*, **560**: A51 DOI: 10.1051/0004-6361/201322551 (see pp. 27, 148)

Adibekyan, V. Zh; Sousa, S. G.; Santos, N. C.; Mena, E. Delgado; Hernández, J. I. González; Israelian, G.; Mayor, M., and Khachatryan, G. (2012b) Chemical Abundances of 1111 FGK Stars from the HARPS GTO Planet Search Program - Galactic Stellar Populations and Planets in: *Astronomy & Astrophysics*, **545**: A32 DOI: 10.1051/0004-6361/201219401 (see pp. 25, 44, 118)

Agol, E. (2011) Transit Surveys for Earths in the Habitable Zones of White Dwarfs in: *The Astrophysical Journal Letters*, **731**:2, L31 DOI: 10.1088/2041-8205/731/2/L31 (see p. 18)

Aigrain, S.; Favata, F., and Gilmore, G. (2004) Characterising Stellar Micro-Variability for Planetary Transit Searches in: *Astronomy & Astrophysics*, **414**:3, 1139–1152 DOI: 10.1051/0004-6361:20034039 (see p. 72)

Aigrain, S.; Llama, J.; Ceillier, T.; Chagas, M. L. das; Davenport, J. R. A.; García, R. A.; Hay, K. L.; Lanza, A. F.; McQuillan, A.; Mazeh, T.; Medeiros, De; R, J.; Nielsen, M. B., and Reinhold, T. (2015) Testing the Recovery of Stellar Rotation Signals from Kepler Light Curves Using a Blind Hare-and-Hounds Exercise in: *Monthly Notices of the Royal Astronomical Society*, **450**:3, 3211–3226 DOI: 10.1093/mnras/stv853 (see pp. 24, 87)

Aigrain, S.; Pont, F., and Zucker, S. (2012) A Simple Method to Estimate Radial Velocity Variations Due to Stellar Activity Using Photometry: Activity-Induced RV Variations from Photometry in: *Monthly Notices of the Royal Astronomical Society*, **419**:4, 3147–3158 DOI: 10.1111/j.1365-2966.2011.19960.x (see p. 90)

Akaike, H. (1973) "Information Theory and an Extension of the Maximum Likelihood Principle" in: *Proceedings of the Second International Symposium on Information Theory* ed. by B. N. Petrov and F. Caski Budapeste: Akademiai Kiado, 267–281 (see p. 49)

Alibert, Y.; Mordasini, C.; Benz, W., and Winisdoerffer, C. (2005) Models of Giant Planet Formation with Migration and Disc Evolution in: *Astronomy and Astrophysics*, **434**: 343–353 DOI: 10.1051/0004-6361:20042032 (see p. 22)

Alonso, R.; Brown, T. M.; Torres, G.; Latham, D. W.; Sozzetti, A.; Georgi Mandushev; Belmonte, J. A.; Charbonneau, D.; Deeg, H. J.; Dunham, E. W.; O'Donovan, F. T., and Stefanik, R. P. (2004) TrES-1: The Transiting Planet of a Bright Ko V Star in: *The Astrophysical Journal Letters*, **613**:2, L153 DOI: 10.1086/425256 (see p. 19)

Ambikasaran, S.; Foreman-Mackey, D.; Greengard, L.; Hogg, D. W., and O'Neil, M. (2014) Fast Direct Methods for Gaussian Processes in: *ArXiv e-prints*, (see p. 88)

Andrae, R.; Schulze-Hartung, T., and Melchior, P. (2010) Dos and Don'ts of Reduced Chi-Squared in: *arXiv:1012.3754 [astro-ph, p.:physics, stat]*, arXiv: 1012.3754 [astro-ph, p.:physics, stat] (see pp. 49, 50)

Anglada-Escudé, G.; Amado, P. J.; Barnes, J.; Berdiñas, Zaira M.; Butler, R. Paul; Coleman, G. A. L.; de la Cueva, Ignacio; Dreizler, S.; Endl, M.; Giesers, B.; Jeffers, S. V.; Jenkins, J. S.; Jones, H. R. A.; Kiraga, M.; Kürster, M.; López-González, M.ía J.; Marvin, C. J.; Morales, N.ás; Morin, J.; Nelson, R. P.; Ortiz, J. L.; Ofir, A.; Paardekooper, S.; Reiners, A.; Rodríguez, E.; Rodríguez-López, C.; Sarmiento, L. F.; Strachan,

J. P.; Tsapras, Y.; Tuomi, M., and Zechmeister, M. (2016a) A Terrestrial Planet Candidate in a Temperate Orbit around Proxima Centauri in: *Nature*, **536**:7617, 437–440 DOI: 10.1038/nature19106 (see p. 18)

Anglada-Escudé, G. and Butler, R. Paul (2012) The HARPS-TERRA Project. I. Description of the Algorithms, Performance, and New Measurements on a Few Remarkable Stars Observed by HARPS in: *The Astrophysical Journal Supplement Series*, **200**:2, 15 DOI: 10.1088/0067-0049/200/2/15 (see p. 43)

Anglada-Escudé, G. and Tuomi, M. (2015) Comment on "Stellar Activity Masquerading as Planets in the Habitable Zone of the M Dwarf Gliese 581" in: *Science*, **347**:6226, 1080–1080 DOI: 10.1126/science.1260796 arXiv: 1503.01976 (see pp. 24, 79)

Anglada-Escudé, G.; Tuomi, M.; Arriagada, P.; Zechmeister, M.; Jenkins, J. S.; Ofir, A.; Dreizler, S.; E. Gerlach; Marvin, C. J.; Reiners, A.; Jeffers, S. V.; Butler, R. Paul; Vogt, S. S.; Amado, P. J.; C. Rodríguez-López; Berdiñas, Z. M.; Morin, J.; Crane, J. D.; Shectman, S. A.; Díaz, M. R.; Sarmiento, L. F., and Jones, H. R. A. (2016b) No Evidence for Activity Correlations in the Radial Velocities of Kapteyn's Star in: *The Astrophysical Journal*, **830**:2, 74 DOI: 10.3847/0004-637X/830/2/74 (see p. 79)

Astudillo-Defru, N.; Bonfils, X.; Delfosse, X.; Ségransan, D.; Forveille, T.; Bouchy, F.; Gillon, M.; Lovis, C.; Mayor, M.; Neves, V.; Pepe, F.; Perrier, C.; Queloz, D.; Rojo, P.; Santos, N. C., and Udry, S. (2015) The HARPS Search for Southern Extra-Solar Planets. XXXVI. Planetary Systems and Stellar Activity of the M Dwarfs GJ 3293, GJ 3341, and GJ 3543 in: *Astronomy and Astrophysics*, **575**: A119 DOI: 10.1051/0004-6361/201424253 (see p. 43)

Backer, D. C.; Foster, R. S., and Sallmen, S. (1993) A Second Companion of the Millisecond Pulsar 1620-26 in: *Nature*, **365**:6449, 817–819 DOI: 10.1038/365817a0 (see p. 18)

Bakos, G. Á; Noyes, R. W.; Kovács, G.; Latham, D. W.; Sasselov, D. D.; Torres, G.; Fischer, D. A.; Stefanik, R. P.; Sato, B.; Johnson, J. A.; Pál, A.; Marcy, G. W.; Butler, R. P.; Esquerdo, G. A.; Stanek, K. Z.; J. Lázár; Papp, I.; Sári, P., and Sipőcz, B. (2007) HAT-P-1b: A Large-Radius, Low-Density Exoplanet Transiting One Member of a Stellar Binary in: *The Astrophysical Journal*, **656**:1, 552 DOI: 10.1086/509874 (see p. 19)

Balan, S. T. and Lahav, O. (2009) EXOFIT: Orbital Parameters of Extrasolar Planets from Radial Velocities in: *Monthly Notices of the Royal Astronomical Society*, **394**:4, 1936–1944 DOI: 10.1111/j.1365-2966.2008.14385.x (see pp. 62, 63)

Baliunas, S. L.; Donahue, R. A.; Soon, W. H.; Horne, J. H.; Frazer, J.; Woodard-Eklund, L.; Bradford, M.; Rao, L. M.; Wilson, O. C.; Zhang, Q.; Bennett, W.; Briggs, J.; Carroll, S. M.; Duncan, D. K.; Figueroa, D.; Lanning, H. H.; Misch, T.; Mueller, J.; Noyes, R. W.; Poppe, D.; Porter, A. C.; Robinson, C. R.; Russell, J.; Shelton, J. C.; Soyumer, T.; Vaughan, A. H., and Whitney, J. H. (1995) Chromospheric Variations in Main-Sequence Stars in: *The Astrophysical Journal*, **438**: 269–287 DOI: 10.1086/175072 (see p. 76)

Baluev, R. V. (2008) Optimal Strategies of Radial Velocity Observations in Planet Search Surveys in: *Monthly Notices of the Royal Astronomical Society*, **389**:3, 1375–1382 DOI: 10.1111/j.1365-2966.2008.13656.x (see p. 149)

– (2013) PlanetPack: A Radial-Velocity Time-Series Analysis Tool Facilitating Exoplanets Detection, Characterization, and Dynamical Simulations in: *Astronomy and Computing*, **2**: 18–26 DOI: 10.1016/j.ascom.2013.07.001 arXiv: 1306.1446 (see p. 37)

Baranne, A.; Queloz, D.; Mayor, M.; Adrianzyk, G.; Knispel, G.; Kohler, D.; Lacroix, D.; Meunier, J.-P.; Rimbaud, G., and Vin, A. (1996) ELODIE: A Spectrograph for Accurate Radial Velocity Measurements. in: *Astronomy and Astrophysics Supplement Series*, **119**: 373–390 (see p. 44)

Barclay, T.; Endl, M.; Huber, D.; Foreman-Mackey, D.; Cochran, W. D.; MacQueen, P. J.; Rowe, J. F., and Quintana, E. V. (2015) Radial Velocity Observations and Light Curve Noise Modeling Confirm That Kepler-91b Is a Giant Planet Orbiting a Giant Star in: *The Astrophysical Journal*, **800**:1, 46 DOI: 10.1088/0004-637X/800/1/46 (see pp. 24, 76)

Barge, P.; Baglin, A.; Auvergne, M.; Rauer, H.; Léger, A.; Schneider, J.; Pont, F.; Aigrain, S.; Almenara, J.-M.; Alonso, R.; Barbieri, M.; Bordé, P.; Bouchy, F.; Deeg, H. J.; la Reza, De; Deleuil, M.; Dvorak, R.; Erikson, A.; Fridlund, M.; Gillon, M.; Gondoin, P.; Guillot, T.; Hatzes, A.; Hebrard, G.; Jorda, L.; Kabath, P.; Lammer, H.; Llebaria, A.; Loeillet, B.; Magain, P.; Mazeh, T.; Moutou, C.; Ollivier, M.; Pätzold, M.; Queloz, D.;

Rouan, D.; Shporer, A., and Wuchterl, G. (2008) Transiting Exoplanets from the CoRoT Space Mission in: *Astronomy and Astrophysics*, **482**:3, L17–L20 DOI: 10.1051/0004-6361:200809353 (see p. 19)

Barnes, J. R.; Jeffers, S. V.; Anglada-Escudé, G.; Haswell, C. A.; Jones, H. R. A.; Tuomi, M.; Feng, F.; Jenkins, J. S., and Petit, P. (2017) Recovering Planet Radial Velocity Signals in the Presence of Starspot Activity in Fully Convective Stars in: *Monthly Notices of the Royal Astronomical Society*, **466**:2, 1733–1740 DOI: 10.1093/mnras/stw3170 (see p. 73)

Barros, S. C. C.; Almenara, J. M.; Deleuil, M.; Díaz, R. F.; Csizmadia, Sz.; Cabrera, J.; Chaintreuil, S.; Collier Cameron, A.; Hatzes, A.; Haywood, R.; Lanza, A. F.; Aigrain, S.; Alonso, R.; Bordé, P.; Bouchy, F.; Deeg, H. J.; Erikson, A.; Fridlund, M.; Grziwa, S.; Gandolfi, D.; Guillot, T.; Guenther, E.; Leger, A.; Moutou, C.; Ollivier, M.; Pasternacki, T.; Pätzold, M.; Rauer, H.; Rouan, D.; Santerne, A.; Schneider, J., and Wuchterl, G. (2014) Revisiting the Transits of CoRoT-7b at a Lower Activity Level in: *Astronomy & Astrophysics*, **569**: A74 DOI: 10.1051/0004-6361/201423939 (see pp. 76, 89)

Batygin, K.; Bodenheimer, P. H., and Laughlin, G. P. (2016) In Situ Formation and Dynamical Evolution of Hot Jupiter Systems in: *The Astrophysical Journal*, **829**:2, 114 DOI: 10.3847/0004-637X/829/2/114 (see p. 22)

Baumann, P.; Ramírez, I.; Meléndez, J.; Asplund, M., and Lind, K. (2010) Lithium Depletion in Solar-like Stars: No Planet Connection in: *Astronomy & Astrophysics*, **519**: A87 DOI: 10.1051/0004-6361/201015137 (see p. 27)

Beaugé, C.; Ferraz-Mello, S., and Michtchenko, T. A. (2012) Multi-Planet Extrasolar Systems — Detection and Dynamics in: *Research in Astronomy and Astrophysics*, **12**:8, 1044 DOI: 10.1088/1674-4527/12/8/009 (see p. 37)

Beaugé, C. and Nesvorný, D. (2013) Emerging Trends in a Period-Radius Distribution of Close-in Planets in: *The Astrophysical Journal*, **763**: 12 DOI: 10.1088/0004-637X/763/1/12 (see p. 27)

Beaulieu, J.-P.; Bennett, D. P.; Fouqué, P.; Williams, A.; Dominik, M.; Jørgensen, U. G.; Kubas, D.; Cassan, A.; Coutures, C.; Greenhill, J.; Hill, K.; Menzies, J.; Sackett, P. D.; Albrow, M.; Brillant, S.; Caldwell, J. a. R.; Calitz, J. J.; Cook, K. H.; Corrales, E.; Desort, M.; Dieters, S.; Dominis, D.; Donatowicz, J.; Hoffman, M.; Kane, S.; Marquette, J.-B.; Martin, R.; Meintjes, P.; Pollard, K.; Sahu, K.; Vinter, C.; Wambsganss, J.; Woller, K.; Horne, K.; Steele, I.; Bramich, D. M.; Burgdorf, M.; Snodgrass, C.; Bode, M.; Udalski, A.; Szymański, M. K.; Kubiak, M.; Wi&ecedil, T.; ckowski; Pietrzyński, G.; Soszyński, I.; Szewczyk, O.; Wyrzykowski, Ł; Paczyński, B.; Abe, F.; Bond, I. A.; Britton, T. R.; Gilmore, A. C.; Hearnshaw, J. B.; Itow, Y.; Kamiya, K.; Kilmartin, P. M.; Korpela, A. V.; Masuda, K.; Matsubara, Y.; Motomura, M.; Muraki, Y.; Nakamura, S.; Okada, C.; Ohnishi, K.; Rattenbury, N. J.; Sako, T.; Sato, S.; Sasaki, M.; Sekiguchi, T.; Sullivan, D. J.; Tristram, P. J.; Yock, P. C. M., and Yoshioka, T. (2006) Discovery of a Cool Planet of 5.5 Earth Masses through Gravitational Microlensing in: *Nature*, **439**:7075, 437–440 DOI: 10.1038/nature04441 (see p. 20)

Bedding, T. R.; Butler, R. Paul; Kjeldsen, H.; Baldry, I. K.; O'Toole, S. J.; Tinney, C.topher G.; Marcy, G. W.; Kienzle, F., and Carrier, F. (2001) Evidence for Solar-like Oscillations in β Hydri in: *The Astrophysical Journal Letters*, **549**:1, L105 DOI: 10.1086/319139 (see p. 74)

Belorizky, D. (1938) Le Soleil, Etoile Variable in: *L'Astronomie*, **52**: 359–361 (see p. 29)

Benedict, G. Fritz; McArthur, B. E.; Gatewood, G.; Nelan, E.; Cochran, W. D.; Artie Hatzes; Endl, M.; Wittenmyer, R.; Baliunas, S. L.; Walker, G. A. H.; Yang, S.; Martin Kürster; Els, S., and Paulson, D. B. (2006) The Extrasolar Planet Eridani b: Orbit and Mass in: *The Astronomical Journal*, **132**:5, 2206 DOI: 10.1086/508323 (see p. 21)

Bennett, D. P.; Bond, I. A.; Udalski, A.; Sumi, T.; Abe, F.; Fukui, A.; Furusawa, K.; Hearnshaw, J. B.; S. Holderness; Itow, Y.; Kamiya, K.; Korpela, A. V.; Kilmartin, P. M.; Lin, W.; Ling, C. H.; Masuda, K.; Y. Matsubara; Miyake, N.; Muraki, Y.; Nagaya, M.; Okumura, T.; Ohnishi, K.; Perrott, Y. C.; Rattenbury, N. J.; T. Sako; Saito, To; Sato, S.; Skuljan, L.; Sullivan, D. J.; Sweatman, W. L.; Tristram, P. J.; Yock, P. C. M.; M. Kubiak; Szymański, M. K.; Pietrzyński, G.; Soszyński, I.; Szewczyk, O.; Wyrzykowski, Ł; Ulaczyk, K.; V. Batista; Beaulieu, J. P.; Brillant, S.; Cassan, A.; Fouqué, P.; Kervella, P.; Kubas, D., and

Marquette, J. B. (2008) A Low-Mass Planet with a Possible Sub-Stellar-Mass Host in Microlensing Event MOA-2007-BLG-192 in: *The Astrophysical Journal*, **684**:1, 663 DOI: 10.1086/589940 (see p. 20)

Berg, B. A. (2004) *Markov Chain Monte Carlo Simulations and Their Statistical Analysis: With Web-Based Fortran Code* OCLC: ocm57519136 Hackensack, NJ: World Scientific (see p. 151)

Berger, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis* 2nd ed Springer series in statistics New York: Springer-Verlag (see p. 51)

Bernardo, J. M. and Smith, A. F. M. (1994) *Bayesian Theory* Wiley series in probability and mathematical statistics Chichester, E.. ; New York: Wiley (see p. 51)

Beuzit, J.; Feldt, M.; Dohlen, K.; Mouillet, D.; Puget, P.; Wildi, F.; Abe, L.; Antichi, J.; Baruffolo, A.; Baudoz, P.; Boccaletti, A.; Carbillet, M.; Charton, J.; Claudi, R.; Downing, M.; Fabron, C.; Feautrier, P.; Fedrigo, E.; Fusco, T.; Gach, J.; Gratton, R.; Henning, T.; Hubin, N.; Joos, F.; Kasper, M.; Langlois, M.; Lenzen, R.; Moutou, C.; Pavlov, A.; Petit, C.; Pragt, J.; Rabou, P.; Rigal, F.; Roelfsema, R.; Rousset, G.; Saisse, M.; Schmid, H.; Stadler, E.; Thalmann, C.; Turatto, M.; Udry, S.; Vakili, F., and Waters, R. (2008) "SPHERE: A 'Planet Finder' Instrument for the VLT" in: vol. 7014, 701418 DOI: 10.1117/12.790120 (see p. 20)

Bodenheimer, P; Hubickyj, O., and Lissauer, J. J. (2000) Models of the in Situ Formation of Detected Extrasolar Giant Planets in: *Icarus*, **143**:1, 2–14 DOI: 10.1006/icar.1999.6246 (see p. 22)

Boisse, I.; Bonfils, X., and Santos, N. C. (2012) SOAP: A Tool for the Fast Computation of Photometry and Radial Velocity Induced by Stellar Spots in: *Astronomy & Astrophysics*, **545**: A109 DOI: 10.1051/0004-6361/201219115 (see p. 85)

Boisse, I.; Bouchy, F.; Hébrard, G.; Bonfils, X.; Santos, N., and Vauclair, S. (2011) Disentangling between Stellar Activity and Planetary Signals in: *Astronomy & Astrophysics*, **528**: A4 DOI: 10.1051/0004-6361/201014354 (see pp. 77, 78, 89)

Boisse, I.; Moutou, C.; Vidal-Madjar, A.; Bouchy, F.; Pont, F.; Hébrard, G.; Bonfils, X.; Croll, B.; Delfosse, X.; Desort, M.; Forveille, T.; Lagrange, A.-M.; Loeillet, B.; Lovis, C.; Matthews, J. M.; Mayor, M.; Pepe, F.; Perrier, C.; Queloz, D.; Rowe, J. F.; Santos, N. C.; Ségransan, D., and Udry, S. (2009) Stellar Activity of Planetary Host Star HD 189 733 in: *Astronomy and Astrophysics*, **495**: 959–966 DOI: 10.1051/0004-6361:200810648 (see pp. 76, 77)

Boley, A. C.; Contreras, A. P. Granados, and Gladman, B. (2016) The In Situ Formation of Giant Planets at Short Orbital Periods in: *The Astrophysical Journal*, **817**:2, L17 DOI: 10.3847/2041-8205/817/2/L17 (see p. 22)

Bond, I. A.; Udalski, A.; Jaroszyñski, M.; Rattenbury, N. J.; Paczyñski, B.; Soszyñski, I.; Wyrzykowski, L.; Szymañski, M. K.; Kubiak, M.; Szewczyk, O.; Żebruñ, K.; Pietrzyñski, G.; Abe, F.; Bennett, D. P.; Eguchi, S.; Furuta, Y.; Hearnshaw, J. B.; Kamiya, K.; Kilmartin, P. M.; Kurata, Y.; Masuda, K.; Matsubara, Y.; Muraki, Y.; Noda, S.; K. Okajima; Sako, T.; Sekiguchi, T.; Sullivan, D. J.; Sumi, T.; Tristram, P. J.; Yanagisawa, T.; Yock, P. C. M.; The MOA, and Collaborations, OGLE (2004) OGLE 2003-BLG-235/MOA 2003-BLG-53: A Planetary Microlensing Event in: *The Astrophysical Journal Letters*, **606**:2, L155 DOI: 10.1086/420928 (see p. 20)

Bonfils, X.; Delfosse, X.; Udry, S.; Forveille, T.; Mayor, M.; Perrier, C.; Bouchy, F.; Gillon, M.; Lovis, C.; Pepe, F.; Queloz, D.; Santos, N. C.; Ségransan, D., and Bertaux, J.-L. (2013) The HARPS Search for Southern Extra-Solar Planets: XXXI. The M-Dwarf Sample in: *Astronomy & Astrophysics*, **549**: A109 DOI: 10.1051/0004-6361/201014704 (see p. 23)

Bonfils, X.; Mayor, M.; Delfosse, X.; Forveille, T.; Gillon, M.; Perrier, C.; Udry, S.; Bouchy, F.; Lovis, C.; Pepe, F.; Queloz, D.; Santos, N. C., and Bertaux, J.-L. (2007) The HARPS Search for Southern Extra-Solar Planets. X. A Msini=11MEarth Planet around the Nearby Spotted M Dwarf GJ674 in: *Astronomy and Astrophysics*, **474**:1, 293–299 DOI: 10.1051/0004-6361:20077068 (see pp. 76, 80, 125)

Bonnefoy, M.; Zurlo, A.; Baudino, J. L.; Lucas, P.; Mesa, D.; Maire, A.-L.; Vigan, A.; Galicher, R.; Homeier, D.; Marocco, F.; Gratton, R.; Chauvin, G.; Allard, F.; Desidera, S.; Kasper, M.; Moutou, C.; Lagrange, A.-M.; Antichi, J.; Baruffolo, A.; Baudrand, J.; Beuzit, J.-L.; Boccaletti, A.; Cantalloube, F.; Carbillet, M.; Charton, J.; Claudi, R. U.; Costille, A.; Dohlen, K.; Dominik, C.; Fantinel, D.; Feautrier, P.; Feldt, M.;

Fusco, T.; Gigan, P.; Girard, J. H.; Gluck, L.; Gry, C.; Henning, T.; Janson, M.; Langlois, M.; Madec, F.; Magnard, Y.; Maurel, D.; Mawet, D.; Meyer, M. R.; Milli, J.; Moeller-Nilsson, O.; Mouillet, D.; Pavlov, A.; Perret, D.; Pujet, P.; Quanz, S. P.; Rochat, S.; Rousset, G.; Roux, A.; Salasnich, B.; Salter, G.; Sauvage, J.-F.; Schmid, H. M.; Sevin, A.; Soenke, C.; Stadler, E.; Turatto, M.; Udry, S.; Vakili, F.; Wahhaj, Z., and Wildi, F. (2016) First Light of the VLT Planet Finder SPHERE - IV. Physical and Chemical Properties of the Planets around HR8799 in: *Astronomy & Astrophysics*, **587**: A58 DOI: 10.1051/0004-6361/201526906 (see p. 20)

Bonomo, A. S. and Lanza, A. F. (2012) Starspot Activity and Rotation of the Planet-Hosting Star Kepler-17 in: *Astronomy & Astrophysics*, **547**: A37 DOI: 10.1051/0004-6361/201219999 (see p. 150)

Borucki, W. J.; Koch, D. G.; Basri, G.; Batalha, N.; Brown, T. M.; Bryson, S. T.; Caldwell, D.; Christensen-Dalsgaard, Jørgen; Cochran, W. D.; DeVore, E.; Dunham, E. W.; III, T. N. Gautier; Geary, J. C.; Gilliland, R.; Gould, A.; Howell, S. B.; Jenkins, J. M.; Latham, D. W.; Lissauer, J. J.; Marcy, G. W.; Rowe, J.; Sasselov, D.; Alan Boss; Charbonneau, D.; Ciardi, D.; Doyle, L.; Dupree, A. K.; Ford, E. B.; Jonathan Fortney; Holman, M. J.; Seager, S.; Steffen, J. H.; Tarter, J.; Welsh, W. F.; Christopher Allen; Buchhave, L. A.; Christiansen, J. L.; Clarke, B. D.; Das, S.; Désert, J.; Michael Endl; Fabrycky, D.; Fressin, F.; Haas, M.; Horch, E.; Howard, A.; Isaacson, H.; Hans Kjeldsen; Kolodziejczak, J.; Kulesa, C.; Li, Jie; Lucas, P. W.; Machalek, P.; Donald McCarthy; MacQueen, P.; Meibom, Søren; Miquel, T.; Prsa, A.; Quinn, S. N.; Quintana, E. V.; Ragozzine, D.; Sherry, W.; Shporer, A.; Tenenbaum, P.; Torres, G.; Twicken, J. D.; Cleve, J. Van; Walkowicz, L.; Witteborn, F. C., and Still, M. (2011) Characteristics of Planetary Candidates Observed by Kepler. II. Analysis of the First Four Months of Data in: *The Astrophysical Journal*, **736**:1, 19 DOI: 10.1088/0004-637X/736/1/19 (see p. 19)

Boss, A. P. (1995) Proximity of Jupiter-Like Planets to Low-Mass Stars in: *Science*, **267**:5196, 360–362 DOI: 10.1126/science.267.5196.360 (see p. 18)

Bouchy, F.; Bazot, M.; Santos, N. C.; Vauclair, S., and Sosnowska, D. (2005) Asteroseismology of the Planet-Hosting Star μ Arae: I. The Acoustic Spectrum in: *Astronomy and Astrophysics*, **440**:2, 609–614 DOI: 10.1051/0004-6361:20052697 (see pp. 74, 98)

Bouchy, F.; Pepe, F., and Queloz, D. (2001) Fundamental Photon Noise Limit to Radial Velocity Measurements in: *Astronomy and Astrophysics*, **374**:2, 733–739 DOI: 10.1051/0004-6361:20010730 (see pp. 40–42)

Bowler, B. P. (2016) Imaging Extrasolar Giant Planets in: *Publications of the Astronomical Society of the Pacific*, **128**:968, 102001 DOI: 10.1088/1538-3873/128/968/102001 (see p. 20)

Bradshaw, S. J. and Hartigan, P. (2014) On Sunspot and Starspot Lifetimes in: *The Astrophysical Journal*, **795**:1, 79 DOI: 10.1088/0004-637X/795/1/79 (see p. 87)

Brewer, B. J. (2014) Inference for Trans-Dimensional Bayesian Models with Diffusive Nested Sampling in: arXiv: 1411.3921 (see pp. 53, 57, 88, 145, 146, 156, 157)

Brewer, B. J. and Donovan, C. P. (2015) Fast Bayesian Inference for Exoplanet Discovery in Radial Velocity Data in: *Monthly Notices of the Royal Astronomical Society*, **448**:4, 3206–3214 DOI: 10.1093/mnras/stv199 (see pp. 28, 52, 55, 62, 145, 146, 157)

Brewer, B. J. and Foreman-Mackey, D. (2016) DNest4: Diffusive Nested Sampling in C++ and Python in: arXiv: 1606.03757 (see p. 157)

Brewer, B. J.; Pártay, L. B., and Csányi, Gábor (2011) Diffusive Nested Sampling in: *Statistics and Computing*, **21**:4, 649–656 DOI: 10.1007/s11222-010-9198-8 (see pp. 53, 145, 154, 155)

Brezinski, C. (2005) La méthode de Cholesky in: *Revue d'histoire des mathématiques*, **11**:2, 205–238 (see p. 88)

Buchhave, L. A. and Latham, D. W. (2015) The Metallicities of Stars with and without Transiting Planets in: *The Astrophysical Journal*, **808**: 187 DOI: 10.1088/0004-637X/808/2/187 (see p. 26)

Buchhave, L. A.; Latham, D. W.; Johansen, A.; Bizzarro, M.; Torres, G.; Rowe, J. F.; Batalha, N. M.; Borucki, W. J.; Brugamyer, E.; Caldwell, C.; Bryson, S. T.; Ciardi, D. R.; Cochran, W. D.; Endl, M.; Esquerdo, G. A.; Ford, E. B.; Geary, J. C.; Gilliland, R. L.; Hansen, T.; Isaacson, H.; Laird, J. B.; Lucas, P. W.; Marcy, G. W.; Morse, J. A.; Robertson, P.; Shporer, A.; Stefanik, R. P.; Still, M., and Quinn, S. N. (2012) An Abundance of

Small Exoplanets around Stars with a Wide Range of Metallicities in: *Nature*, DOI: 10.1038/nature11121 (see pp. 26, 148)

Bumba, V and Kleczek, J (1976) *Basic Mechanisms of Solar Activity* OCLC: 840307916 Dordrecht: Springer Netherlands (see p. 150)

Burgasser, A. J.; Kirkpatrick, J. Davy; Reid, I. Neill; Brown, M. E.; Miskey, C. L., and Gizis, J. E. (2003) Binarity in Brown Dwarfs: T Dwarf Binaries Discovered with the Hubble Space Telescope Wide Field Planetary Camera 2 in: *The Astrophysical Journal*, **586**:1, 512 DOI: 10.1086/346263 (see p. 141)

Burnham, K. P. and Anderson, D. R. (2010) *Model Selection and Multimodel Inference* 2nd Springer (see p. 50)

Butler, R. P.; Wright, J. T.; Marcy, G. W.; Fischer, D. A.; Vogt, S. S.; Tinney, C. G.; Jones, H. R. A.; Carter, B. D.; Johnson, J. A.; McCarthy, C., and Penny, A. J. (2006) Catalog of Nearby Exoplanets in: *The Astrophysical Journal*, **646**:1, 505–522 DOI: 10.1086/504701 (see p. 17)

Butler, R. Paul and Marcy, G. W. (1996) A Planet Orbiting 47 Ursae Majoris in: *The Astrophysical Journal Letters*, **464**:2, L153 (see p. 18)

Butler, R. Paul; Marcy, G. W.; Fischer, D. A.; Brown, T. M.; Contos, A. R.; Korzennik, S. G.; Nisenson, P., and Noyes, R. W. (1999) Evidence for Multiple Companions to υ Andromedae in: *The Astrophysical Journal*, **526**:2, 916 DOI: 10.1086/308035 (see p. 18)

Butler, R. Paul; Marcy, G. W.; Vogt, S. S.; Fischer, D. A.; Henry, G. W.; Gregory Laughlin, and Wright, J. T. (2003) Seven New Keck Planets Orbiting G and K Dwarfs in: *The Astrophysical Journal*, **582**:1, 455 DOI: 10.1086/344570 (see p. 65)

Butler, R. Paul; Marcy, G. W.; Williams, E.; Hauser, H., and Shirts, P. (1997) Three New "51 Pegasi–Type" Planets in: *The Astrophysical Journal*, **474**:2, L115–L118 DOI: 10.1086/310444 (see p. 18)

Butler, R. Paul; Vogt, S. S.; Marcy, G. W.; Fischer, D. A.; Wright, J. T.; Henry, G. W.; Laughlin, G., and Lissauer, J. J. (2004) A Neptune-Mass Planet Orbiting the Nearby M Dwarf GJ 436 in: *The Astrophysical Journal*, **617**:1, 580 DOI: 10.1086/425173 (see p. 18)

Calderhead, B. and Girolami, M. (2009) Estimating Bayes Factors via Thermodynamic Integration and Population MCMC in: *Computational Statistics & Data Analysis*, **53**:12, 4028–4045 DOI: 10.1016/j.csda.2009.07.025 (see p. 151)

Cameron, A. Collier (2012) Extrasolar Planets: Astrophysical False Positives in: *Nature*, **492**:7427, 48–50 DOI: 10.1038/492048a (see p. 19)

Campbell, B.; Walker, G. A. H., and Yang, S. (1988) A Search for Substellar Companions to Solar-Type Stars in: *The Astrophysical Journal*, **331**: 902 DOI: 10.1086/166608 (see p. 17)

Carolo, E.; Desidera, S.; Gratton, R.; Fiorenzano, A. F. Martinez; Marzari, F.; Endl, M.; Mesa, D.; Barbieri, M.; Cecconi, M.; Claudi, R. U.; Cosentino, R., and Scuderi, S. (2014) A Vigorous Activity Cycle Mimicking a Planetary System in HD 200466 in: *Astronomy & Astrophysics*, **567**: A48 DOI: 10.1051/0004-6361/201323102 (see pp. 76, 77)

Carter, J. A.; Agol, E.; Chaplin, W. J.; Basu, S.; Bedding, T. R.; Buchhave, L. A.; Christensen-Dalsgaard, Jørgen; Deck, K. M.; Elsworth, Y.; Fabrycky, D. C.; Ford, E. B.; Fortney, J. J.; Hale, S. J.; Handberg, R.; Hekker, S.; Holman, M. J.; Huber, D.; Karoff, C.; Kawaler, S. D.; Kjeldsen, H.; Lissauer, J. J.; Lopez, E. D.; Lund, M. N.; Lundkvist, M.; Metcalfe, T. S.; Miglio, A.; Rogers, L. A.; Stello, D.; Borucki, W. J.; Bryson, S.; Christiansen, J. L.; Cochran, W. D.; Geary, J. C.; Gilliland, R. L.; Haas, M. R.; Hall, J.; Howard, A. W.; Jenkins, J. M.; Klaus, T.; Koch, D. G.; Latham, D. W.; MacQueen, P. J.; Sasselov, D.; Steffen, J. H.; Twicken, J. D., and Winn, J. N. (2012) Kepler-36: A Pair of Planets with Neighboring Orbits and Dissimilar Densities in: *Science*, **337**:6094, 556–559 DOI: 10.1126/science.1223269 (see p. 19)

Cassan, A.; Kubas, D.; Beaulieu, J.-P.; Dominik, M.; Horne, K.; Greenhill, J.; Wambsganss, J.; Menzies, J.; Williams, A.; Jørgensen, U. G.; Udalski, A.; Bennett, D. P.; Albrow, M. D.; Batista, V.; Brillant, S.; Caldwell, J. a. R.; Cole, A.; Coutures, Ch; Cook, K. H.; Dieters, S.; Prester, D. Dominis; Donatowicz, J.; Fouqué, P.; Hill, K.; Kains, N.; Kane, S.; Marquette, J.-B.; Martin, R.; Pollard, K. R.; Sahu, K. C.; Vinter, C.; Warren, D.; Watson, B.; Zub, M.; Sumi, T.; Szymañski, M. K.; Kubiak, M.; Poleski, R.; Soszynski, I.; Ulaczyk, K.;

Pietrzyński, G., and Wyrzykowski, Ł (2012) One or More Bound Planets per Milky Way Star from Microlensing Observations in: *Nature*, **481**:7380, 167–169 DOI: 10.1038/nature10684 (see p. 20)

Cegla, H. M.; Watson, C. A.; Marsh, T. R.; Shelyag, S.; Moulds, V.; Littlefair, S.; Mathioudakis, M.; Pollacco, D., and Bonfils, X. (2012) Stellar Jitter from Variable Gravitational Redshift: Implications for Radial Velocity Confirmation of Habitable Exoplanets in: *Monthly Notices of the Royal Astronomical Society*, **421**: L54–L58 DOI: 10.1111/j.1745-3933.2011.01205.x (see p. 40)

Chabrier, G. and Küker, M. (2006) Large-Scale $\vec \alpha\mathsf{â}$-Dynamo in Low-Mass Stars and Brown Dwarfs in: *Astronomy & Astrophysics*, **446**:3, 1027–1037 DOI: 10.1051/0004-6361:20042475 (see p. 72)

Charbonneau, D.; Brown, T. M.; Latham, D. W., and Mayor, M. (2000) Detection of Planetary Transits Across a Sun-like Star in: *The Astrophysical Journal Letters*, **529**:1, L45 DOI: 10.1086/312457 (see pp. 18, 20)

Charbonneau, P. (2010) Dynamo Models of the Solar Cycle in: *Living Reviews in Solar Physics*, **7**:1, 3 DOI: 10.12942/lrsp-2010-3 (see pp. 71, 76)

Chauvin, G.; Lagrange, A.-M.; Dumas, C.; Zuckerman, B.; Mouillet, D.; Song, I.; Beuzit, J.-L., and Lowrance, P. (2004) A Giant Planet Candidate near a Young Brown Dwarf - Direct VLT/NACO Observations Using IR Wavefront Sensing in: *Astronomy & Astrophysics*, **425**:2, L29–L32 DOI: 10.1051/0004-6361:200400056 (see p. 20)

Chib, S. and Jeliazkov, I. (2001) Marginal Likelihood From the Metropolis–Hastings Output in: *Journal of the American Statistical Association*, **96**:453, 270–281 DOI: 10.1198/016214501750332848 (see p. 151)

Cholesky, A. (1910) *Sur La Résolution Numérique Des Systèmes d'équations Linéaires* (see p. 88)

Christensen-Dalsgaard, JØrgen (2004) Physics of Solar-like Oscillations in: *Solar Physics*, **220**:2, 137–168 DOI: 10.1023/B:SOLA.0000031392.43227.7d (see p. 74)

Cincunegui, C.; Díaz, R. F., and Mauas, P. J. D. (2007) Hα and the Ca II H and K Lines as Activity Proxies for Late-Type Stars in: *Astronomy and Astrophysics*, **469**: DOI: 10.1051/0004-6361:20066503 (see p. 79)

Clanton, C. and Gaudi, B. Scott (2016) Synthesizing Exoplanet Demographics: A Single Population Of Long-Period Planetary Companions To M Dwarfs Consistent With Microlensing, Radial Velocity, And Direct Imaging Surveys in: *The Astrophysical Journal*, **819**:2, 125 DOI: 10.3847/0004-637X/819/2/125 (see p. 18)

Claudi, R. (2016) "Direct Imaging of Faint Companions" in: *Methods of Detecting Exoplanets* ed. by V. Bozza; L. Mancini, and A. Sozzetti vol. 428 Cham: Springer International Publishing, 183–252 DOI: 10.1007/978-3-319-27458-4_4 (see p. 20)

Cochran, W. D. and Endl, M. (2008) Using the Hobby–Eberly Telescope to Place Constraints on Planetary System Formation in: *Physica Scripta*, **2008**:T130, 014006 DOI: 10.1088/0031-8949/2008/T130/014006 (see p. 97)

Cochran, W. D.; Endl, M.; Wittenmyer, R. A., and Bean, J. L. (2007) A Planetary System around HD 155358: The Lowest Metallicity Planet Host Star in: *The Astrophysical Journal*, **665**:2, 1407 DOI: 10.1086/519555 (see p. 97)

Cochran, W. D.; Fabrycky, D. C.; Torres, G.; Fressin, F.; Désert, J.; Darin Ragozzine; Sasselov, D.; Fortney, J. J.; Rowe, J. F.; Brugamyer, E. J.; Bryson, S. T.; Carter, J. A.; Ciardi, D. R.; Howell, S. B.; Steffen, J. H.; Borucki, W. J.; Koch, D. G.; Winn, J. N.; Welsh, W. F.; Uddin, K.; Tenenbaum, P.; Still, M.; Seager, S.; Quinn, S. N.; Mullally, F.; Miller, N.; Marcy, G. W.; MacQueen, P. J.; Lucas, P.; Lissauer, J. J.; Latham, D. W.; Knutson, H.; Kinemuchi, K.; Johnson, J. A.; Jenkins, J. M.; Howard Isaacson; Howard, A.; Horch, E.; Holman, M. J.; Henze, C. E.; Haas, M. R.; Gilliland, R. L.; III, T. N. Gautier; Ford, E. B.; Fischer, D. A.; Everett, M.; Michael Endl; Demory, B.; Deming, D.; Charbonneau, D.; Caldwell, D.; Buchhave, L.; Brown, T. M., and Batalha, N. (2011) Kepler-18b, c, and d: A System of Three Planets Confirmed by Transit Timing Variations, Light Curve Validation, Warm-Spitzer Photometry, and Radial Velocity Measurements in: *The Astrophysical Journal Supplement Series*, **197**:1, 7 DOI: 10.1088/0067-0049/197/1/7 (see p. 38)

Cochran, W. D.; Hatzes, A. P., and Hancock, T. J. (1991) Constraints on the Companion Object to HD 114762 in: *The Astrophysical Journal*, **380**: L35–L38 DOI: 10.1086/186167 (see p. 17)

Correia, A. C. M.; Couetdic, J.; Laskar, J.; Bonfils, X.; Mayor, M.; Bertaux, J.-L.; Bouchy, F.; Delfosse, X.; Forveille, T.; Lovis, C.; Pepe, F.; Perrier, C.; Queloz, D., and Udry, S. (2010) The HARPS Search for Southern Extra-Solar Planets: XIX. Characterization and Dynamics of the GJ 876 Planetary System in: *Astronomy and Astrophysics*, **511**: A21 DOI: 10.1051/0004-6361/200912700 (see pp. 23, 38)

Correia, A. C. M.; Udry, S.; Mayor, M.; Benz, W.; Bertaux, J.-L.; Bouchy, F.; Laskar, J.; Lovis, C.; Mordasini, C.; Pepe, F., and Queloz, D. (2009) The HARPS Search for Southern Extra-Solar Planets. XVI. HD 45364, a Pair of Planets in a 3:2 Mean Motion Resonance in: *Astronomy and Astrophysics*, **496**:2, 521–526 DOI: 10.1051/0004-6361:200810774 (see p. 23)

Correia, A. C. M.; Udry, S.; Mayor, M.; Laskar, J.; Naef, D.; Pepe, F.; Queloz, D., and Santos, N. C. (2005) The CORALIE Survey for Southern Extra-Solar Planets. XIII. A Pair of Planets around HD 202206 or a Circumbinary Planet? in: *Astronomy and Astrophysics*, **440**: 751–758 DOI: 10.1051/0004-6361:20042376 (see p. 39)

Couetdic, J.; Laskar, J.; Correia, A. C. M.; Mayor, M., and Udry, S. (2010) Dynamical Stability Analysis of the HD 202206 System and Constraints to the Planetary Orbits in: *Astronomy & Astrophysics*, **519**: A10 DOI: 10.1051/0004-6361/200913635 (see p. 39)

Courcol, B.; Bouchy, F., and Deleuil, M. (2016) An Upper Boundary in the Mass-Metallicity Plane of Exo-Neptunes in: *Monthly Notices of the Royal Astronomical Society*, **461**:2, 1841–1849 DOI: 10.1093/mnras/stw1049 (see pp. 26, 148, 149)

Cover, T. M. and Thomas, J. A. (2006) *Elements of Information Theory* 2nd ed Hoboken, N.J: Wiley-Interscience (see pp. 54, 60)

Cumming, A. (2004) Detectability of Extrasolar Planets in Radial Velocity Surveys in: *Monthly Notices of the Royal Astronomical Society*, **354**:4, 1165–1176 DOI: 10.1111/j.1365-2966.2004.08275.x (see p. 139)

Cumming, A.; Marcy, G. W., and Butler, R. Paul (1999) The Lick Planet Search: Detectability and Mass Thresholds in: *The Astrophysical Journal*, **526**:2, 890 (see pp. 139, 140)

Cunha, M. S. (2009) Magnetoacoustic Oscillations in Ap Stars in: *Proceedings of the International Astronomical Union*, **5**:H15, 362–363 DOI: 10.1017/S1743921310009816 (see p. 72)

da Silva, L.; Girardi, L.; Pasquini, L.; Setiawan, J.; Lühe, O. von der; Medeiros, J. R. de; Hatzes, A.; P. Döllinger, M., and Weiss, A. (2006) Basic Physical Parameters of a Selected Sample of Evolved Stars in: *Astronomy & Astrophysics*, **458**:2, 609–623 DOI: 10.1051/0004-6361:20065105 (see p. 101)

Da Silva, R.; Udry, S.; Bouchy, F.; Mayor, M.; Moutou, C.; Pont, F.; Queloz, D.; Santos, N. C.; Ségransan, D., and Zucker, S. (2006) Elodie Metallicity-Biased Search for Transiting Hot Jupiters in: *Astronomy and Astrophysics*, **446**:2, 717–722 DOI: 10.1051/0004-6361:20054116 (see p. 97)

Dawson, R. I.; Chiang, E., and Lee, E. J. (2015) A Metallicity Recipe for Rocky Planets in: *Monthly Notices of the Royal Astronomical Society*, **453**:2, 1471–1483 DOI: 10.1093/mnras/stv1639 (see p. 26)

Dawson, R. I. and Fabrycky, D. C. (2010) Radial Velocity Planets De-Aliased: A New, Short Period For Super-Earth 55 Cnc E in: *The Astrophysical Journal*, **722**:1, 937–953 DOI: 10.1088/0004-637X/722/1/937 (see p. 141)

Delfosse, X.; Forveille, T.; Mayor, M.; Perrier, C.; Naef, D., and Queloz, D. (1998) The Closest Extrasolar Planet. A Giant Planet around the M4 Dwarf GL 876 in: *Astronomy and Astrophysics*, **338**: L67–L70 (see p. 18)

Desidera, S.; Gratton, R. G.; Endl, M.; Claudi, R. U.; Cosentino, R.; Barbieri, M.; Bonanno, G.; Lucatello, S.; Martinez Fiorenzano, A. F.; Marzari, F., and Scuderi, S. (2004) No Planet around HD 219542 B in: *Astronomy and Astrophysics*, **420**:3, L27–L30 DOI: 10.1051/0004-6361:20040155 (see p. 76)

Desort, M.; Lagrange, A.-M.; Galland, F.; Udry, S., and Mayor, M. (2007) Search for Exoplanets with the Radial-Velocity Technique: Quantitative Diagnostics of Stellar Activity in: *Astronomy and Astrophysics*, **473**: 983–993 DOI: 10.1051/0004-6361:20078144 (see pp. 76, 77)

Díaz, R. F.; Ségransan, D.; Udry, S.; Lovis, C.; Pepe, F.; Dumusque, X.; Marmier, M.; Alonso, R.; Benz, W.; Bouchy, F.; Coffinet, A.; Cameron, A. Collier; Deleuil, M.; Figueira, P.; Gillon, M.; Curto, G. Lo; Mayor, M.; Mordasini, C.; Motalebi, F.; Moutou, C.; Pollacco, D.; Pompei, E.; Queloz, D.; Santos, N., and Wyttenbach,

A. (2016) The HARPS Search for Southern Extra-Solar Planets - XXXVIII. Bayesian Re-Analysis of Three Systems. New Super-Earths, Unconfirmed Signals, and Magnetic Cycles in: *Astronomy & Astrophysics*, **585**: A134 DOI: 10.1051/0004-6361/201526729 (see pp. 76, 80)

Donati, J.-F.; Semel, M.; Carter, B. D.; Rees, D. E., and Collier Cameron, A. (1997) Spectropolarimetric Observations of Active Stars in: *Monthly Notices of the Royal Astronomical Society*, **291**: 658 DOI: 10.1093/mnras/291.4.658 (see p. 73)

Dong, S.; Zheng, Z.; Zhu, Z.; Cat, P. De; Fu, J. N.; Yang, X. H.; Zhang, H.; Jin, Ge, and Zhang, Y. (2014) On the Metallicities of Kepler Stars in: *The Astrophysical Journal Letters*, **789**:1, L3 DOI: 10.1088/2041-8205/789/1/L3 (see p. 26)

Doppler, C. (1842) "Über das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels" Prag (see p. 40)

Dorn, C.; Khan, A.; Heng, K.; Alibert, Y.; Connolly, J. A. D.; Benz, W., and Tackley, P. (2015) Can We Constrain Interior Structure of Rocky Exoplanets from Mass and Radius Measurements? in: *Astronomy & Astrophysics*, DOI: 10.1051/0004-6361/201424915 arXiv: 1502.03605 (see p. 27)

Dorn, C.; Venturini, J.; Khan, A.; Heng, K.; Alibert, Y.; Helled, R.; Rivoldini, A., and Benz, W. (2017) A Generalized Bayesian Inference Method for Constraining the Interiors of Super Earths and Sub-Neptunes in: *Astronomy & Astrophysics*, **597**: A37 DOI: 10.1051/0004-6361/201628708 (see p. 19)

Dravins, D. (1982) Photospheric Spectrum Line Asymmetries and Wavelength Shifts in: *Annual Review of Astronomy and Astrophysics*, **20**: 61–89 DOI: 10.1146/annurev.aa.20.090182.000425 (see p. 73)

Dravins, D.; Lindegren, L., and Madsen, Søren (1999) Astrometric Radial Velocities. I. Non-Spectroscopic Methods for Measuring Stellar Radial Velocity in: *Astronomy and Astrophysics*, **348**: (see p. 40)

Dumusque, X. (2016) Radial Velocity Fitting Challenge. I. Simulating the Data Set Including Realistic Stellar Radial-Velocity Signals in: *Astronomy & Astrophysics*, DOI: 10.1051/0004-6361/201628672 (see p. 79)

Dumusque, X.; Boisse, I., and Santos, N. C. (2014a) SOAP 2.0: A Tool to Estimate the Photometric and Radial Velocity Variations Induced by Stellar Spots and Plages in: *The Astrophysical Journal*, **796**: 132 DOI: 10.1088/0004-637X/796/2/132 (see pp. 73, 74, 85, 150)

Dumusque, X.; Bonomo, A. S.; Haywood, R. D.; Malavolta, L.; Ségransan, D.; Buchhave, L. A.; Collier Cameron, A.; Latham, D. W.; Molinari, E.; Pepe, F.; Udry, S.; Charbonneau, D.; Cosentino, R.; Dressing, C. D.; Figueira, P.; Fiorenzano, A. F. M.; Gettel, S.; Harutyunyan, A.; Horne, K.; Lopez-Morales, M.; Lovis, C.; Mayor, M.; Micela, G.; Motalebi, F.; Nascimbeni, V.; Phillips, D. F.; Piotto, G.; Pollacco, D.; Queloz, D.; Rice, K.; Sasselov, D.; Sozzetti, A.; Szentgyorgyi, A., and Watson, C. A. (2014b) The Kepler-10 Planetary System Revisited by HARPS-N: A Hot Rocky World and a Solid Neptune-Mass Planet in: *The Astrophysical Journal*, **789**: 154 DOI: 10.1088/0004-637X/789/2/154 (see p. 19)

Dumusque, X.; Borsa, F.; Damasso, M.; Díaz, R. F.; Gregory, P. C.; Hara, N. C.; Hatzes, A.; Rajpaul, V.; Tuomi, M.; Aigrain, S.; Anglada-Escudé, G.; Bonomo, A. S.; Boué, G.; Dauvergne, F.; Frustagli, G.; Giacobbe, P.; Haywood, R. D.; Jones, H. R. A.; Laskar, J.; Pinamonti, M.; Poretti, E.; Rainer, M.; Ségransan, D.; Sozzetti, A., and Udry, S. (2017) Radial-Velocity Fitting Challenge - II. First Results of the Analysis of the Data Set in: *Astronomy & Astrophysics*, **598**: A133 DOI: 10.1051/0004-6361/201628671 (see pp. 125, 146)

Dumusque, X.; Glenday, A.; Phillips, D. F.; Buchschacher, N.; Collier Cameron, A.; Cecconi, M.; Charbonneau, D.; Cosentino, R.; Ghedina, A.; Latham, D. W.; Li, Chih-Hao; Lodi, M.; Lovis, C.; Molinari, E.; Pepe, F.; Udry, S.; Sasselov, D.; Szentgyorgyi, A., and Walsworth, R. (2015a) HARPS-N Observes the Sun as a Star in: *The Astrophysical Journal Letters*, **814**: L21 DOI: 10.1088/2041-8205/814/2/L21 (see p. 150)

Dumusque, X.; Lovis, C.; Ségransan, D.; Mayor, M.; Udry, S.; Benz, W.; Bouchy, F.; Lo Curto, G.; Mordasini, C.; Pepe, F.; Queloz, D.; Santos, N. C., and Naef, D. (2011a) The HARPS Search for Southern Extra-Solar Planets: XXX. Planetary Systems around Stars with Solar-like Magnetic Cycles and Short-Term Activity Variation⋆ ⋆ ⋆ in: *Astronomy & Astrophysics*, **535**: A55 DOI: 10.1051/0004-6361/201117148 (see p. 76)

Dumusque, X.; Pepe, F.; Lovis, C., and Latham, D. W. (2015b) Characterization of a Spurious One-Year Signal in HARPS Data in: *The Astrophysical Journal*, **808**:2, 171 DOI: 10.1088/0004-637X/808/2/171 (see pp. 127, 149)

Dumusque, X.; Pepe, F.; Lovis, C.; Ségransan, D.; Sahlmann, J.; Benz, W.; Bouchy, F.; Mayor, M.; Queloz, D.; Santos, N., and Udry, S. (2012) An Earth-Mass Planet Orbiting α Centauri B in: *Nature*, **491**:7423, 207–211 DOI: 10.1038/nature11572 (see p. 139)

Dumusque, X.; Santos, N. C.; Udry, S.; Lovis, C., and Bonfils, X. (2011b) Planetary Detection Limits Taking into Account Stellar Noise: II. Effect of Stellar Spot Groups on Radial-Velocities in: *Astronomy & Astrophysics*, **527**: A82 DOI: 10.1051/0004-6361/201015877 (see pp. 24, 76, 77)

Dumusque, X.; Udry, S.; Lovis, C.; Santos, N. C., and Monteiro, M. J. P. F. G. (2011c) Planetary Detection Limits Taking into Account Stellar Noise: I. Observational Strategies to Reduce Stellar Oscillation and Granulation Effects⋆ in: *Astronomy & Astrophysics*, **525**: A140 DOI: 10.1051/0004-6361/201014097 (see pp. 24, 72, 74, 76, 77, 98)

Eastman, J.; Gaudi, B. Scott, and Agol, E. (2013) EXOFAST: A Fast Exoplanetary Fitting Suite in IDL in: *Publications of the Astronomical Society of the Pacific*, **125**:923, 83 DOI: 10.1086/669497 (see p. 56)

Efron, B. and Gous, A. (2001) Scales of Evidence for Model Selection: Fisher versus Jeffreys in: *Lecture Notes-Monograph Series*, **38**: ed. by R. E. Kass; G. S. Datta, and P. Lahiri, 208–256 (see p. 60)

Einstein, A. (1905) Zur Elektrodynamik bewegter Körper in: *Annalen der Physik*, **322**:10, 891–921 DOI: 10.1002/andp.19053221004 (see p. 40)

Endl, M.; Kürster, M.; Els, S.; Hatzes, A. P., and Cochran, W. D. (2001) The Planet Search Program at the ESO Coudé Echelle Spectrometer. II. The Alpha Centauri System: Limits for Planetary Companions in: *Astronomy and Astrophysics*, **374**: 675–681 DOI: 10.1051/0004-6361:20010723 (see p. 139)

Ercolano, B. and Clarke, C. J. (2010) Metallicity, Planet Formation and Disc Lifetimes in: *Monthly Notices of the Royal Astronomical Society*, **402**:4, 2735–2743 DOI: 10.1111/j.1365-2966.2009.16094.x (see pp. 27, 148)

Fabrycky, D. C. (2010) "Non-Keplerian Dynamics" in: *Exoplanets*, arXiv:1006.3834 (see p. 38)

Fabrycky, D. C.; Ford, E. B.; Steffen, J. H.; Rowe, J. F.; Carter, J. A.; Moorhead, A. V.; Batalha, N. M.; Borucki, W. J.; Bryson, S.; Buchhave, L. A.; Christiansen, J. L.; Ciardi, D. R.; Cochran, W. D.; Endl, M.; Fanelli, M. N.; Debra Fischer; Fressin, F.; Geary, J.; Haas, M. R.; Hall, J. R.; Holman, M. J.; Jenkins, J. M.; Koch, D. G.; Latham, D. W.; Li, Jie; Lissauer, J. J.; Lucas, P.; Marcy, G. W.; Tsevi Mazeh; McCauliff, S.; Quinn, S.; Ragozzine, D.; Sasselov, D., and Shporer, A. (2012) Transit Timing Observations from Kepler. IV. Confirmation of Four Multiple-Planet Systems by Simple Physical Models in: *The Astrophysical Journal*, **750**:2, 114 DOI: 10.1088/0004-637X/750/2/114 (see p. 39)

Faria, J. P.; Haywood, R. D.; Brewer, B. J.; Figueira, P.; Oshagh, M.; Santerne, A., and Santos, N. C. (2016a) Uncovering the Planets and Stellar Activity of CoRoT-7 Using Only Radial Velocities in: *Astronomy & Astrophysics*, **588**: A31 DOI: 10.1051/0004-6361/201527899 (see pp. 28, 61, 71, 80, 87, 93, 147, 157, 158)

Faria, J. P.; Santos, N. C.; Figueira, P.; Mortier, A.; Dumusque, X.; Boisse, I.; Lo Curto, G.; Lovis, C.; Mayor, M.; Melo, C.; Pepe, F.; Queloz, D.; Santerne, A.; Ségransan, D.; Sousa, S. G.; Sozzetti, A., and Udry, S. (2016b) The HARPS Search for Southern Extra-Solar Planets: XL. Searching for Neptunes around Metal-Poor Stars⋆ in: *Astronomy & Astrophysics*, **589**: A25 DOI: 10.1051/0004-6361/201527522 (see pp. 28, 139, 141, 142)

Feng, F.; Tuomi, M., and Jones, H. R. A. (2017) Agatha: Disentangling Periodic Signals from Correlated Noise in a Periodogram Framework in: *Monthly Notices of the Royal Astronomical Society*, **470**:4, 4794–4814 DOI: 10.1093/mnras/stx1126 (see p. 125)

Feroz, F.; Balan, S. T., and Hobson, M. P. (2011a) Bayesian Evidence for Two Companions Orbiting HIP 5158 in: *Monthly Notices of the Royal Astronomical Society*, **416**: L104–L108 DOI: 10.1111/j.1745-3933.2011.01109.x (see p. 154)

– (2011b) Detecting Extrasolar Planets from Stellar Radial Velocities Using Bayesian Evidence in: *Monthly Notices of the Royal Astronomical Society*, **415**:4, 3462–3472 (see pp. 66, 145, 146, 154)

Feroz, F. and Hobson, M. P. (2008) Multimodal Nested Sampling: An Efficient and Robust Alternative to Markov Chain Monte Carlo Methods for Astronomical Data Analyses: Multimodal Nested Sampling in:

Monthly Notices of the Royal Astronomical Society, **384**:2, 449–463 DOI: 10.1111/j.1365-2966.2007.12353.x (see p. 154)

– (2014) Bayesian Analysis of Radial Velocity Data of GJ667C with Correlated Noise: Evidence for Only Two Planets in: *Monthly Notices of the Royal Astronomical Society*, **437**:4, 3540–3549 DOI: 10.1093/mnras/stt2148 (see pp. 103, 154)

Feroz, F.; Hobson, M. P., and Bridges, M. (2009) MultiNest: An Efficient and Robust Bayesian Inference Tool for Cosmology and Particle Physics in: *Monthly Notices of the Royal Astronomical Society*, **398**:4, 1601–1614 (see p. 154)

Feroz, F.; Hobson, M. P.; Cameron, E., and Pettitt, A. N. (2013) Importance Nested Sampling and the MultiNest Algorithm in: arXiv: 1306.2144 (see p. 154)

Ferraz-Mello, S.; Tadeu dos Santos, M.; Beaugé, C.; Michtchenko, T. A., and Rodríguez, A. (2011) On the Mass Determination of Super-Earths Orbiting Active Stars: The CoRoT-7 System in: *Astronomy & Astrophysics*, **531**: A161 DOI: 10.1051/0004-6361/201016059 (see p. 89)

Figueira, P.; Adibekyan, V. Zh; Oshagh, M.; Neal, J. J.; Rojas-Ayala, B.; Lovis, C.; Melo, C.; Pepe, F.; Santos, N. C., and Tsantaki, M. (2016) Radial Velocity Information Content of M Dwarf Spectra in the Near-Infrared in: *Astronomy & Astrophysics*, **586**: A101 DOI: 10.1051/0004-6361/201526900 (see p. 43)

Figueira, P.; Marmier, M.; Bonfils, X.; di Folco, E.; Udry, S.; Santos, N. C.; Lovis, C.; Mégevand, D.; Melo, C. H. F.; Pepe, F.; Queloz, D.; Ségransan, D.; Triaud, A. H. M. J., and Viana Almeida, P. (2010) Evidence against the Young Hot-Jupiter around BD +20 1790 in: *Astronomy and Astrophysics*, **513**: L8 DOI: 10.1051/0004-6361/201014323 (see pp. 24, 76, 125)

Figueira, P.; Santos, N. C.; Pepe, F.; Lovis, C., and Nardetto, N. (2013) Line-Profile Variations in Radial-Velocity Measurements: Two Alternative Indicators for Planetary Searches in: *Astronomy & Astrophysics*, **557**: A93 DOI: 10.1051/0004-6361/201220779 (see pp. 77, 78)

– (2015) Line-Profile Variations in Radial-Velocity Measurements (Corrigendum) - Two Alternative Indicators for Planetary Searches in: *Astronomy & Astrophysics*, **582**: C2 DOI: 10.1051/0004-6361/201220779e (see p. 78)

Fischer, D. A.; Anglada-Escude, G.; Arriagada, P.; Baluev, R. V.; Bean, J. L.; Bouchy, F.; Buchhave, L. A.; Carroll, T.; Chakraborty, A.; Crepp, J. R.; Dawson, R. I.; Diddams, S. A.; Dumusque, X.; Eastman, J. D.; Endl, M.; Figueira, P.; Ford, E. B.; Daniel Foreman-Mackey; Fournier, P.; Fürész, G.; Gaudi, B. Scott; Gregory, P. C.; Grundahl, F.; Hatzes, A. P.; Hébrard, G.; Herrero, E.; Hogg, D. W.; Howard, A. W.; Johnson, J. A.; Paul Jorden; Jurgenson, C. A.; Latham, D. W.; Laughlin, G.; Loredo, T. J.; Lovis, C.; Suvrath Mahadevan; McCracken, T. M.; Pepe, F.; Perez, M.; Phillips, D. F.; Plavchan, P. P.; Lisa Prato; Quirrenbach, A.; Reiners, A.; Robertson, P.; Santos, N. C.; Sawyer, D.; Damien Segransan; Sozzetti, A.; Steinmetz, T.; Szentgyorgyi, A.; Udry, S.; Valenti, J. A.; Wang, S. X.; Wittenmyer, R. A., and Wright, J. T. (2016) State of the Field: Extreme Precision Radial Velocities in: *Publications of the Astronomical Society of the Pacific*, **128**:964, 066001 DOI: 10.1088/1538-3873/128/964/066001 (see pp. 44, 76, 145, 146, 149)

Fischer, D. A.; Laughlin, G.; Butler, P.; Marcy, G.; Johnson, J.; Henry, G.; Valenti, J.; Steve Vogt; Ammons, M.; Robinson, S.; Spear, G.; Strader, J.; Driscoll, P.; Fuller, A.; Teresa Johnson; Manrao, E.; McCarthy, C.; Muñoz, Melesio; Tah, K. L.; Wright, J.; Ida, S.; Bun'ei Sato; Toyota, E., and Minniti, D. (2005) The N2K Consortium. I. A Hot Saturn Planet Orbiting HD 88133 in: *The Astrophysical Journal*, **620**:1, 481 DOI: 10.1086/426810 (see p. 97)

Fischer, D. A. and Valenti, J. (2005) The Planet-Metallicity Correlation in: *The Astrophysical Journal*, **622**:2, 1102 (see p. 25)

Ford, E. B. (2003) *Selected Topics on Extrasolar Planetary Systems*. PhD thesis Princeton University (see p. 53)

– (2005) Quantifying the Uncertainty in the Orbits of Extrasolar Planets in: *The Astronomical Journal*, **129**:3, 1706 (see pp. 37, 53, 145)

– (2008) Adaptive Scheduling Algorithms for Planet Searches in: *The Astronomical Journal*, **135**:3, 1008 DOI: 10.1088/0004-6256/135/3/1008 (see p. 149)

Ford, E. B. (2014) Architectures of Planetary Systems and Implications for Their Formation in: *Proceedings of the National Academy of Sciences*, **111**:35, 12616–12621 DOI: 10.1073/pnas.1304219111 (see p. 23)

Ford, E. B. and Gregory, P. C. (2007) "Bayesian Model Selection and Extrasolar Planet Detection" in: *Statistical Challenges in Modern Astronomy IV* ed. by G. J. Babu and E. D. Feigelson vol. 371 Astronomical Society of the Pacific Conference Series, 189 (see pp. 51, 56)

Foreman-Mackey, D.; Agol, E.; Ambikasaran, S., and Angus, R. (2017) Fast and Scalable Gaussian Process Modeling with Applications to Astronomical Time Series in: *The Astronomical Journal*, **154**:6, 220 DOI: 10.3847/1538-3881/aa9332 (see p. 88)

Forveille, T.; Bonfils, X.; Delfosse, X.; Gillon, M.; Udry, S.; Bouchy, F.; Lovis, C.; Mayor, M.; Pepe, F.; Perrier, C.; Queloz, D.; Santos, N., and Bertaux, J.-L. (2009) The HARPS Search for Southern Extra-Solar Planets. XIV. Gl 176b, a Super-Earth Rather than a Neptune, and at a Different Period in: *Astronomy and Astrophysics*, **493**:2, 645–650 DOI: 10.1051/0004-6361:200810557 (see p. 24)

Gaia Collaboration et al. (2016) Gaia Data Release 1. Summary of the Astrometric, Photometric, and Survey Properties in: *Astronomy and Astrophysics*, **595**: A2 DOI: 10.1051/0004-6361/201629512 (see p. 185)

Gamerman, D. and Lopes, H. Freitas (2006) *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* 2nd ed Texts in statistical science series 68 Boca Raton: Taylor & Francis (see p. 151)

Gaudi, B. Scott (2012) Microlensing Surveys for Exoplanets in: *Annual Review of Astronomy and Astrophysics*, **50**:1, 411–453 DOI: 10.1146/annurev-astro-081811-125518 (see p. 20)

Geyer, C. J. (1992) Practical Markov Chain Monte Carlo in: *Statistical Science*, **7**:4, 473–483 DOI: 10.1214/ss/1177011137 (see p. 151)

Giguere, M. J.; Fischer, D. A.; Zhang, C. X. Y.; Matthews, J. M.; Cameron, C., and Henry, G. W. (2016) A Combined Spectroscopic and Photometric Stellar Activity Study of Epsilon Eridani in: *The Astrophysical Journal*, **824**:2, 150 DOI: 10.3847/0004-637X/824/2/150 (see pp. 76, 125)

Gomes da Silva, J.; Santos, N. C.; Bonfils, X.; Delfosse, X.; Forveille, T., and Udry, S. (2011) Long-Term Magnetic Activity of a Sample of M-Dwarf Stars from the HARPS Program. I. Comparison of Activity Indices in: *Astronomy and Astrophysics*, **534**: A30 DOI: 10.1051/0004-6361/201116971 (see p. 79)

González Hernández, J. I.; Israelian, G.; Santos, N. C.; Sousa, S.; Delgado-Mena, E.; Neves, V., and Udry, S. (2010) Searching for the Signatures of Terrestrial Planets in Solar Analogs in: *The Astrophysical Journal*, **720**:2, 1592–1602 DOI: 10.1088/0004-637X/720/2/1592 (see p. 27)

Gonzalez, G. (1997) The Stellar Metallicity-Giant Planet Connection in: *Monthly Notices of the Royal Astronomical Society*, **285**: 403–412 (see p. 25)

Gould, A. (2016) "Microlensing Planets" in: *Methods of Detecting Exoplanets* ed. by V. Bozza; L. Mancini, and A. Sozzetti vol. 428 Cham: Springer International Publishing, 135–179 DOI: 10.1007/978-3-319-27458-4_3 (see p. 20)

Goździewski, K.; Konacki, M., and Maciejewski, A. J. (2006) Orbital Configurations and Dynamical Stability of Multiplanet Systems around Sun-like Stars HD 202206, 14 Herculis, HD 37124, and HD 108874 in: *The Astrophysical Journal*, **645**:1, 688 DOI: 10.1086/504030 (see pp. 39, 136)

Gray, D. F. (1989) The Morphology of Reversed Spectral Line Bisectors in: *Publications of the Astronomical Society of the Pacific*, **101**:643, 832 DOI: 10.1086/132503 (see p. 77)

– (2005) *The Observation and Analysis of Stellar Photospheres* 3rd OCLC: 254321006 Cambridge Univ. Press (see pp. 41–43, 100)

– (2009) The Third Signature of Stellar Granulation in: *The Astrophysical Journal*, **697**:2, 1032 DOI: 10.1088/0004-637X/697/2/1032 (see pp. 73, 74)

Gray, R. O. and Corbally, C. J. (2009) *Stellar Spectral Classification* Princeton series in astrophysics OCLC: ocn276340686 Princeton, N.J: Princeton University Press (see p. 41)

Green, P. J. (1995) Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination in: *Biometrika*, **82**:4, 711 DOI: 10.2307/2337340 (see p. 53)

Gregory, P. C. (2005) A Bayesian Analysis of Extrasolar Planet Data for HD 73526 in: *The Astrophysical Journal*, **631**:2, 1198 (see pp. 53, 145)

– (2010) *Bayesian Logical Data Analysis for the Physical Sciences* Cambridge University Press (see p. 53)
– (2011) Bayesian Exoplanet Tests of a New Method for MCMC Sampling in Highly Correlated Model Parameter Spaces in: *Monthly Notices of the Royal Astronomical Society*, **410**:1, 94–110 DOI: 10.1111/j.1365-2966.2010.17428.x (see p. 151)
– (2012) Additional Keplerian Signals in the HARPS Data for Gliese 667C from a Bayesian Re-Analysis in: *arXiv:1212.4058*, arXiv: 1212.4058 (see p. 55)
– (2016) An Apodized Kepler Periodogram for Separating Planetary and Stellar Activity Signals in: *Monthly Notices of the Royal Astronomical Society*, **458**:3, 2604–2633 DOI: 10.1093/mnras/stw147 (see p. 57)
Grether, D. and Lineweaver, C. H. (2006) How Dry Is the Brown Dwarf Desert? Quantifying the Relative Number of Planets, Brown Dwarfs, and Stellar Companions around Nearby Sun-like Stars in: *The Astrophysical Journal*, **640**:2, 1051–1062 DOI: 10.1086/500161 (see p. 23)
Grieves, N.; Ge, Jian; Thomas, N.; Ma, Bo; Sithajan, S.; Ghezzi, L.; Kimock, B.; Willis, K.; De Lee, N.; Lee, B.; Fleming, S. W.; Agol, E.; Troup, N.; Paegert, M.; Schneider, D. P.; Stassun, K.; Varosi, F.; Zhao, Bo; Jian, L.; Li, Rui; Mello, P. de; F, Gustavo; Bizyaev, D.; Pan, K.; Dutra-Ferreira, L.ícia; Lorenzo-Oliveira, D.; Santiago, B.ílio X.; Costa, Da; N, Luiz; Maia, M. A. G.; Ogando, R. L. C.; Peloso, D., and F, E. (2017) Exploring the Brown Dwarf Desert: New Substellar Companions from the SDSS-III MARVELS Survey in: *Monthly Notices of the Royal Astronomical Society*, **467**:4, 4264–4281 DOI: 10.1093/mnras/stx334 (see p. 23)
Grunblatt, S. K.; Howard, A. W., and Haywood, R. D. (2015) Determining the Mass of Kepler-78b with Nonparametric Gaussian Process Estimation in: *The Astrophysical Journal*, **808**:2, 127 DOI: 10.1088/0004-637X/808/2/127 (see pp. 85, 147)
Hall, J. C. (2008) Stellar Chromospheric Activity in: *Living Reviews in Solar Physics*, **5**: DOI: 10.12942/lrsp-2008-2 (see pp. 79, 80)
Hall, J. C.; Henry, G. W.; Lockwood, G. Wesley; Skiff, B. A., and Saar, S. H. (2009) The Activity and Variability of the Sun and Sun-Like Stars. II. Contemporaneous Photometry and Spectroscopy of Bright Solar Analogs in: *The Astronomical Journal*, **138**:1, 312 DOI: 10.1088/0004-6256/138/1/312 (see p. 150)
Han, E.; Wang, S. X.; Wright, J. T.; Feng, Y. Katherina; Zhao, M.; Fakhouri, O.; Brown, J. I., and Hancock, C. (2014) Exoplanet Orbit Database. II. Updates to Exoplanets.Org in: *Publications of the Astronomical Society of the Pacific*, **126**:943, 827 DOI: 10.1086/678447 (see p. 17)
Hanslmeier, A.; Nesis, A., and Mattig, W. (1991) The Variation of the Solar Granulation Structure in Active and Non-Active Regions in: *Astronomy and Astrophysics*, **251**: 307–311 (see p. 74)
Harvey, J. (1985) "High-Resolution Helioseismology" in: vol. 235 (see p. 72)
Hasegawa, Y. and Hirashita, H. (2014) Planet Traps And First Planets: The Critical Metallicity For Gas Giant Formation in: *The Astrophysical Journal*, **788**:1, 62 DOI: 10.1088/0004-637X/788/1/62 (see pp. 25, 148, 150)
Hasegawa, Y. and Pudritz, R. E. (2014) Planet Traps and Planetary Cores: Origins of the Planet-Metallicity Correlation in: *The Astrophysical Journal*, **794**:1, 25 DOI: 10.1088/0004-637X/794/1/25 (see pp. 25, 26, 148)
Hathaway, D. H. (2015) The Solar Cycle in: *arXiv:1502.07020 [astro-ph]*, arXiv: 1502.07020 [astro-ph] (see pp. 76, 150)
Hatzes, A. P. (2014) The Detection of Earth-Mass Planets around Active Stars. The Mass of Kepler-78b in: *Astronomy and Astrophysics*, **568**: A84 DOI: 10.1051/0004-6361/201424025 (see p. 76)
Hatzes, A. P. and Cochran, W. D. (1992) "Spectrograph Requirements for Precise Radial Velocity Measurements" in: vol. 40, 275 (see p. 40)
– (1993) Long-Period Radial Velocity Variations in Three K Giants in: *The Astrophysical Journal*, **413**: 339–348 DOI: 10.1086/173002 (see p. 17)
Hatzes, A. P.; Cochran, W. D.; Endl, M.; Guenther, E. W.; Saar, S. H.; Walker, G. a. H.; Yang, S.; Hartmann, M.; Esposito, M.; Paulson, D. B., and Döllinger, M. P. (2006) Confirmation of the Planet Hypothesis for the Long-Period Radial Velocity Variations of β Geminorum in: *Astronomy & Astrophysics*, **457**:1, 335–341 DOI: 10.1051/0004-6361:20065445 (see p. 18)

Hatzes, A. P.; Cochran, W. D.; Endl, M.; McArthur, B.; Paulson, D. B.; Walker, G. A. H.; Campbell, B., and Yang, S. (2003) A Planetary Companion to γ Cephei A in: *The Astrophysical Journal*, **599**:2, 1383 DOI: 10.1086/379281 (see p. 17)

Hatzes, A. P.; Dvorak, R.; Wuchterl, G.; Guterman, P.; Hartmann, M.; Fridlund, M.; Gandolfi, D.; Guenther, E., and Pätzold, M. (2010) An Investigation into the Radial Velocity Variations of CoRoT-7 in: *Astronomy and Astrophysics*, **520**: A93 DOI: 10.1051/0004-6361/201014795 (see pp. 76, 89, 91, 147)

Hatzes, A. P.; Fridlund, M.; Nachmani, G.; Mazeh, T.; Valencia, D.; Hébrard, G.; Carone, L.; Pätzold, M.; Udry, S.; Bouchy, F.; Deleuil, M.; Moutou, C.; Barge, P.; Bordé, P.; Deeg, H.; Tingley, B.; Dvorak, R.; Gandolfi, D.; Ferraz-Mello, S.; Wuchterl, Günther; Guenther, E.; Guillot, T.; Rauer, H.; Erikson, A.; Cabrera, J.; Csizmadia, S.; Léger, A.; Lammer, H.; Weingrill, Jörg; Queloz, D.; Alonso, R.; Rouan, D., and Schneider, J. (2011) The Mass of CoRoT-7b in: *The Astrophysical Journal*, **743**:1, 75 DOI: 10.1088/0004-637X/743/1/75 (see p. 89)

Haywood, R. D.; Collier Cameron, A.; Queloz, D.; Barros, S. C. C.; Deleuil, M.; Fares, R.; Gillon, M.; Lanza, A. F.; Lovis, C.; Moutou, C.; Pepe, F.; Pollacco, D.; Santerne, A.; Ségransan, D., and Unruh, Y. C. (2014) Planets and Stellar Activity: Hide and Seek in the CoRoT-7 System in: *Monthly Notices of the Royal Astronomical Society*, **443**: 2517–2531 DOI: 10.1093/mnras/stu1320 (see pp. 76, 85, 89, 90, 92, 95, 147)

Heintz, W. D. (1978) Reexamination of Suspected Unresolved Binaries in: *The Astrophysical Journal*, **220**: 931 DOI: 10.1086/155982 (see p. 21)

Henry, G. W.; Marcy, G. W.; Butler, R. Paul, and Vogt, S. S. (2000) A Transiting "51 Peg-like" Planet in: *The Astrophysical Journal Letters*, **529**:1, L41 DOI: 10.1086/312458 (see pp. 18, 20)

Hessman, F. V.; Dhillon, V. S.; Winget, D. E.; Schreiber, M. R.; Horne, K.; Marsh, T. R.; Guenther, E.; Schwope, A., and Heber, U. (2010) On the Naming Convention Used for Multiple Star Systems and Extrasolar Planets in: arXiv: 1012.0707 (see p. 18)

Ho, Shirley and Turner, E. L. (2011) The Posterior Distribution of Sin(i) Values for Exoplanets with MT Sin(i) Determined from Radial Velocity Data in: *The Astrophysical Journal*, **739**:1, 26 DOI: 10.1088/0004-637X/739/1/26 (see p. 38)

Holczer, T.; Mazeh, T.; Nachmani, G.; Jontof-Hutter, D.; Ford, E. B.; Fabrycky, D.; Darin Ragozzine; Kane, M., and Steffen, J. H. (2016) Transit Timing Observations from Kepler. IX. Catalog of the Full Long-Cadence Data Set in: *The Astrophysical Journal Supplement Series*, **225**:1, 9 DOI: 10.3847/0067-0049/225/1/9 (see p. 19)

Holman, M. J.; Fabrycky, D. C.; Ragozzine, D.; Ford, E. B.; Steffen, J. H.; Welsh, W. F.; Lissauer, J. J.; Latham, D. W.; Marcy, G. W.; Walkowicz, L. M.; Batalha, N. M.; Jenkins, J. M.; Rowe, J. F.; Cochran, W. D.; Fressin, F.; Torres, G.; Buchhave, L. A.; Sasselov, D. D.; Borucki, W. J.; Koch, D. G.; Basri, G.; Brown, T. M.; Caldwell, D. A.; Charbonneau, D.; Dunham, E. W.; Gautier, T. N.; Geary, J. C.; Gilliland, R. L.; Haas, M. R.; Howell, S. B.; Ciardi, D. R.; Endl, M.; Fischer, D.; Fürész, Gábor; Hartman, J. D.; Isaacson, H.; Johnson, J. A.; MacQueen, P. J.; Moorhead, A. V.; Morehead, R. C., and Orosz, J. A. (2010) Kepler-9: A System of Multiple Planets Transiting a Sun-Like Star, Confirmed by Timing Variations in: *Science*, **330**:6000, 51–54 DOI: 10.1126/science.1195778 (see p. 19)

Holman, M. J. and Murray, N. W. (2005) The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets in: *Science*, **307**:5713, 1288–1291 DOI: 10.1126/science.1107822 (see p. 19)

Holman, M. J.; Winn, J. N.; Latham, D. W.; O'Donovan, F. T.; Charbonneau, D.; Bakos, Gáspár A.; Esquerdo, G. A.; Hergenrother, C.; Everett, M. E., and Pál, A. (2006) The Transit Light Curve Project. I. Four Consecutive Transits of the Exoplanet XO-1b in: *The Astrophysical Journal*, **652**:2, 1715 DOI: 10.1086/508155 (see p. 19)

Houdebine, E. R.; Mullan, D. J.; Bercu, B.; Paletou, F., and Gebran, M. (2017) The Rotation–Activity Correlations in K and M Dwarfs. II. New Constraints on the Dynamo Mechanisms in Late-K and M Dwarfs Before and At the Transition to Complete Convection in: *The Astrophysical Journal*, **837**:1, 96 DOI: 10.3847/1538-4357/aa5cad (see p. 72)

Howard, A. W.; Marcy, G. W.; Johnson, J. Asher; Fischer, D. A.; Wright, J. T.; Isaacson, H.; Valenti, J. A.; Anderson, J.; Lin, D. N. C., and Ida, S. (2010) The Occurrence and Mass Distribution of Close-in Super-Earths, Neptunes, and Jupiters in: *Science*, **330**:6004, 653–655 DOI: 10.1126/science.1194854 (see p. 23)

Huélamo, N.; Figueira, P.; Bonfils, X.; Santos, N. C.; Pepe, F.; Gillon, M.; Azevedo, R.; Barman, T.; Fernández, M.; di Folco, E.; Guenther, E. W.; Lovis, C.; Melo, C. H. F.; Queloz, D., and Udry, S. (2008) TW Hydrae: Evidence of Stellar Spots Instead of a Hot Jupiter in: *Astronomy and Astrophysics*, **489**:2, L9–L13 DOI: 10.1051/0004-6361:200810596 (see p. 125)

Hussain, G.a.j. (2002) Starspot Lifetimes in: *Astronomische Nachrichten*, **323**:3-4, 349–356 DOI: 10.1002/1521-3994(200208)323:3/4<349::AID-ASNA349>3.0.CO;2-E (see p. 87)

Ida, S. and Lin, D. N. C. (2004) Toward a Deterministic Model of Planetary Formation. II. The Formation and Retention of Gas Giant Planets around Stars with a Range of Metallicities in: *The Astrophysical Journal*, **616**: 567–572 DOI: 10.1086/424830 (see pp. 23, 25, 148)

Iglesias-Marzoa, R.ón; López-Morales, M., and Morales, M.ía Jesús Arévalo (2015) The Rvfit Code: A Detailed Adaptive Simulated Annealing Code for Fitting Binaries and Exoplanets Radial Velocities in: *Publications of the Astronomical Society of the Pacific*, **127**:952, 567 DOI: 10.1086/682056 (see p. 47)

Israelian, G.; Delgado Mena, E.; Santos, N. C.; Sousa, S. G.; Mayor, M.; Udry, S.; Domínguez Cerdeña, Carolina; Rebolo, R., and Randich, S. (2009) Enhanced Lithium Depletion in Sun-like Stars with Orbiting Planets in: *Nature*, **462**: DOI: 10.1038/nature08483 (see p. 27)

Jaynes, E. T. (2003) *Probability Theory: The Logic of Science* Cambridge, UK ; New York, NY: Cambridge University Press (see pp. 51, 54)

Jeffreys, H. (1961) *Theory of Probability* 3rd Third edition Oxford: Clarendon Press (see pp. 51, 59, 60, 146)

Jenkins, J. S. and Tuomi, M. (2014) The Curious Case of HD 41248. A Pair of Static Signals Buried Behind Red Noise in: *The Astrophysical Journal*, **794**:2, 110 (see pp. 24, 122, 124, 125, 147)

Jenkins, J. S.; Tuomi, M.; Brasser, R.; Ivanyuk, O., and Murgas, F. (2013) Two Super-Earths Orbiting the Solar Analog HD 41248 on the Edge of a 7:5 Mean Motion Resonance in: *The Astrophysical Journal*, **771**: 41 DOI: 10.1088/0004-637X/771/1/41 (see pp. 98, 122, 123, 147)

Johnson, J. Asher; Aller, K. M.; Howard, A. W., and Crepp, J. R. (2010) Giant Planet Occurrence in the Stellar Mass-Metallicity Plane in: *Publications of the Astronomical Society of the Pacific*, **122**:894, 905–915 DOI: 10.1086/655775 (see p. 25)

Johnson, J. Asher; Payne, M.; Howard, A. W.; Clubb, K. I.; Ford, E. B.; Bowler, B. P.; Henry, G. W.; Fischer, D. A.; Marcy, G. W.; Brewer, J. M.; Schwab, C.; Sabine Reffert, and Lowe, T. B. (2011) Retired A Stars and Their Companions. VI. A Pair of Interacting Exoplanet Pairs Around the Subgiants 24 Sextanis and HD 200964 in: *The Astronomical Journal*, **141**:1, 16 DOI: 10.1088/0004-6256/141/1/16 (see p. 39)

Johnson, M. C.; Endl, M.; Cochran, W. D.; Meschiari, S.; Robertson, P.; MacQueen, P. J.; Brugamyer, E. J.; Caldwell, C.; Hatzes, A. P.; Ramírez, I., and Wittenmyer, R. A. (2016) A 12-Year Activity Cycle for HD 219134 in: *arXiv:1602.05200 [astro-ph]*, arXiv: 1602.05200 [astro-ph] (see pp. 76, 80, 137, 139)

Jones, E.; Oliphant, T.; Peterson, P., et al. (2001) *SciPy: Open Source Scientific Tools for Python* [Online; accessed 2016-11-29] (see p. 50)

Kalas, P.; Graham, J. R.; Chiang, E.; Fitzgerald, M. P.; Clampin, M.; Kite, E. S.; Stapelfeldt, K.; Marois, C., and Krist, J. (2008) Optical Images of an Exosolar Planet 25 Light-Years from Earth in: *Science*, **322**:5906, 1345–1348 DOI: 10.1126/science.1166609 (see p. 20)

Kallinger, T.; Ridder, J. De; Hekker, S.; Mathur, S.; Mosser, B.; Gruberbauer, M.; García, R. A.; Karoff, C., and Ballot, J. (2014) The Connection between Stellar Granulation and Oscillation as Seen by the Kepler Mission in: *Astronomy & Astrophysics*, **570**: A41 DOI: 10.1051/0004-6361/201424313 (see pp. 72, 74)

Kane, S. R.; Thirumalachari, B.; Henry, G. W.; Hinkel, N. R.; Jensen, E. L. N.; Boyajian, T. S.; Fischer, D. A.; Howard, A. W.; Isaacson, H. T., and Wright, J. T. (2016) Stellar Activity and Exclusion of the Outer Planet in the HD 99492 System in: *The Astrophysical Journal Letters*, **820**:1, L5 DOI: 10.3847/2041-8205/820/1/L5 (see p. 76)

Karoff, C. (2008) *Observational Asteroseismology*. PhD thesis (see p. 72)

Kass, R. E. and Raftery, A. E. (1995) Bayes Factors in: *Journal of the American Statistical Association*, **90**:430, 773 DOI: 10.2307/2291091 (see pp. 59, 60)

Kepler, J. (1609) *Astronomia Nova* Heidelberg: G. Voegelinus (see pp. 31, 32)

– (1619) *Harmonices Mundi Libri V* Austria: Linz (see p. 32)

Kipping, D. M. (2013) Parametrizing the Exoplanet Eccentricity Distribution with the Beta Distribution in: *Monthly Notices of the Royal Astronomical Society: Letters*, **434**:1, L51–L55 DOI: 10.1093/mnrasl/slt075 (see p. 57)

Kjeldsen, H.; Bedding, T. R.; Butler, R. Paul; Christensen-Dalsgaard, Jørgen; Kiss, L. L.; Chris McCarthy; Marcy, G. W.; Tinney, C.topher G., and Wright, J. T. (2005) Solar-like Oscillations in α Centauri B in: *The Astrophysical Journal*, **635**:2, 1281 DOI: 10.1086/497530 (see pp. 72, 74)

Kokubo, E. and Ida, S. (2002) Formation of Protoplanet Systems and Diversity of Planetary Systems in: *The Astrophysical Journal*, **581**:1, 666 DOI: 10.1086/344105 (see p. 25)

Konstantopoulos, I. S. (2015) The Starfish Diagram: Visualising Data within the Context of Survey Samples in: *Astronomy and Computing*, **10**: 116–120 DOI: 10.1016/j.ascom.2015.01.007 (see p. 102)

Kopp, G. and Lean, J. L. (2011) A New, Lower Value of Total Solar Irradiance: Evidence and Climate Significance in: *Geophysical Research Letters*, **38**:1, n/a–n/a DOI: 10.1029/2010GL045777 (see p. 89)

Kullback, S. and Leibler, R. A. (1951) On Information and Sufficiency in: *The Annals of Mathematical Statistics*, **22**:1, 79–86 DOI: 10.1214/aoms/1177729694 (see p. 60)

Lagrange, A.-M.; Bonnefoy, M.; Chauvin, G.; Apai, D.; Ehrenreich, D.; Boccaletti, A.; Gratadour, D.; Rouan, D.; Mouillet, D.; Lacour, S., and Kasper, M. (2010) A Giant Planet Imaged in the Disk of the Young Star β Pictoris in: *Science*, **329**:5987, 57–59 DOI: 10.1126/science.1187187 (see p. 20)

Lagrange, A.-M.; Desort, M.; Galland, F.; Udry, S., and Mayor, M. (2009) Extrasolar Planets and Brown Dwarfs around A–F Type Stars - VI. High Precision RV Survey of Early Type Dwarfs with HARPS in: *Astronomy & Astrophysics*, **495**:1, 335–352 DOI: 10.1051/0004-6361:200810105 (see p. 139)

Lanza, A. F.; Bonomo, A. S.; Moutou, C.; Pagano, I.; Messina, S.; Leto, G.; Cutispoto, G.; Aigrain, S.; Alonso, R.; Barge, P.; Deleuil, M.; Auvergne, M.; Baglin, A., and Collier Cameron, A. (2010) Photospheric Activity, R., and Radial Velocity Variations of the Planet-Hosting Star CoRoT-7 in: *Astronomy and Astrophysics*, **520**: A53 DOI: 10.1051/0004-6361/201014403 (see pp. 89, 92)

Lartillot, N. and Philippe, H. (2006) Computing Bayes Factors Using Thermodynamic Integration in: *Systematic Biology*, **55**:2, 195–207 DOI: 10.1080/10635150500433722 (see p. 151)

Latham, D. W.; Mazeh, T.; Stefanik, R. P.; Mayor, M., and Burki, G. (1989) The Unseen Companion of HD114762: A Probable Brown Dwarf in: *Nature*, **339**:6219, 38–40 DOI: 10.1038/339038a0 (see p. 17)

Lee, M. D. and Wagenmakers, E. (2013) *Bayesian Cognitive Modeling: A Practical Course* OCLC: ocn861318341 Cambridge ; New York: Cambridge University Press (see p. 59)

Léger, A. et al. (2009) Transiting Exoplanets from the CoRoT Space Mission. VIII. CoRoT-7b: The First Super-Earth with Measured Radius in: *Astronomy and Astrophysics*, **506**: 287–302 DOI: 10.1051/0004-6361/200911933 (see pp. 88, 89, 92)

Leighton, R. B.; Noyes, R. W., and Simon, G. W. (1962) Velocity Fields in the Solar Atmosphere. I. Preliminary Report. in: *The Astrophysical Journal*, **135**: 474 DOI: 10.1086/147285 (see p. 74)

Levenberg, K. (1944) A Method for the Solution of Certain Non-Linear Problems in Least Squares in: *Quarterly of Applied Mathematics*, **2**:2, 164–168 (see p. 46)

Lindegren, L. and Dravins, D. (2003) The Fundamental Definition of "Radial Velocity" in: *Astronomy & Astrophysics*, **401**:3, 1185–1201 DOI: 10.1051/0004-6361:20030181 (see pp. 40, 73)

Lindegren, L.; Lammers, U.; Bastian, U.; Hernández, J.; Klioner, S.; Hobbs, D.; Bombrun, A.; Michalik, D.; Ramos-Lerate, M.; Butkevich, A.; Comoretto, G.; Joliet, E.; Holl, B.; Hutton, A.; Parsons, P.; Steidelmüller, H.; Abbas, U.; Altmann, M.; Andrei, A.; Anton, S.; Bach, N.; Barache, C.; Becciani, U.; Berthier, J.; Bianchi, L.; Biermann, M.; Bouquillon, S.; Bourda, G.; Brüsemeister, T.; Bucciarelli, B.; Busonero, D.; Carlucci, T.; Castañeda, J.; Charlot, P.; Clotet, M.; Crosta, M.; Davidson, M.; de Felice, F.; Drimmel, R.; Fabricius, C.;

Fienga, A.; Figueras, F.; Fraile, E.; Gai, M.; Garralda, N.; Geyer, R.; González-Vidal, J. J.; Guerra, R.; Hambly, N. C.; Hauser, M.; Jordan, S.; Lattanzi, M. G.; Lenhardt, H.; Liao, S.; Löffler, W.; McMillan, P. J.; Mignard, F.; Mora, A.; Morbidelli, R.; Portell, J.; Riva, A.; Sarasso, M.; Serraller, I.; Siddiqui, H.; Smart, R.; Spagna, A.; Stampa, U.; Steele, I.; Taris, F.; Torra, J.; van Reeven, W.; Vecchiato, A.; Zschocke, S.; de Bruijne, J.; Gracia, G.; Raison, F.; Lister, T.; Marchant, J.; Messineo, R.; Soffel, M.; Osorio, J.; de Torres, A., and O'Mullane, W. (2016) Gaia Data Release 1: Astrometry: One Billion Positions, Two Million Proper Motions and Parallaxes in: *Astronomy & Astrophysics*, **595**: A4 DOI: 10.1051/0004-6361/201628714 (see pp. 108, 185)

Lissauer, J. J.; Dawson, R. I., and Tremaine, S. (2014) Advances in Exoplanet Science from Kepler in: *Nature*, **513**:7518, 336–344 DOI: 10.1038/nature13781 (see p. 23)

Lissauer, J. J.; Hubickyj, O.; D'Angelo, G., and Bodenheimer, P. (2009) Models of Jupiter's Growth Incorporating Thermal and Hydrodynamic Constraints in: *Icarus*, **199**:2, 338–350 DOI: 10.1016/j.icarus.2008.10.004 (see p. 25)

Lo Curto, G.; Pepe, F.; Avila, G.; Boffin, H.; Bovay, S.; Chazelas, B.; Coffinet, A.; Fleury, M.; Hughes, I.; Lovis, C.; Maire, C.; Manescau, A.; Pasquini, L.; Rihs, S.; Sinclaire, P., and Udry, S. (2015) HARPS Gets New Fibres After 12 Years of Operations in: *The Messenger*, **162**: 9–15 (see pp. 45, 104, 107)

Lockwood, G. W.; Skiff, B. A.; Henry, G. W.; Henry, S.; Radick, R. R.; Baliunas, S. L.; Donahue, R. A., and Soon, W. (2007) Patterns of Photometric and Chromospheric Variation among Sun-like Stars: A 20 Year Perspective in: *The Astrophysical Journal Supplement Series*, **171**:1, 260 DOI: 10.1086/516752 (see p. 80)

Lomb, N. R. (1976) Least-Squares Frequency Analysis of Unequally Spaced Data in: *Astrophysics and Space Science*, **39**:2, 447–462 DOI: 10.1007/BF00648343 (see pp. 47, 162)

Lopez, S. and Jenkins, J. S. (2012) The Effects of Viewing Angle on the Mass Distribution of Exoplanets in: *The Astrophysical Journal*, **756**:2, 177 DOI: 10.1088/0004-637X/756/2/177 (see p. 38)

López-Morales, M.; Haywood, R. D.; Coughlin, J. L.; Zeng, Li; Buchhave, L. A.; Giles, H. A. C.; Affer, L.; Bonomo, A. S.; Charbonneau, D.; Cameron, A. Collier; Rosario Consentino; Dressing, C. D.; Dumusque, X.; Figueira, P.; Fiorenzano, A. F. M.; Avet Harutyunyan; Johnson, J. Asher; Latham, D. W.; Lopez, E. D.; Lovis, C.; Malavolta, L.; Michel Mayor; Micela, G.; Molinari, E.; Mortier, A.; Motalebi, F.; Nascimbeni, V.; Francesco Pepe; Phillips, D. F.; Piotto, G.; Pollacco, D.; Queloz, D.; Rice, K.; Sasselov, D.; Damien Segransan; Sozzetti, A.; Udry, S.; Vanderburg, A., and Watson, C. A. (2016) Kepler-21b: A Rocky Planet Around a V = 8.25 Magnitude Star in: *The Astronomical Journal*, **152**:6, 204 DOI: 10.3847/0004-6256/152/6/204 (see pp. 19, 146)

Loredo, T. J. (1994) "The Return of the Prodigal: Bayesian Inference in Astrophysics" in: Valencia (see p. 51)

– (2004) "Bayesian Adaptive Exploration" in: vol. 707 AIP Conf. Proc., 330–346 DOI: 10.1063/1.1751377 (see p. 149)

Lovis, C.; Dumusque, X.; Santos, N. C.; Bouchy, F.; Mayor, M.; Pepe, F.; Queloz, D.; Ségransan, D., and Udry, S. (2011a) The HARPS Search for Southern Extra-Solar Planets. XXXI. Magnetic Activity Cycles in Solar-Type Stars: Statistics and Impact on Precise Radial Velocities in: arXiv: 1107.5325 (see pp. 24, 76, 79, 137, 139)

Lovis, C.; Mayor, M.; Bouchy, F.; Pepe, F.; Queloz, D.; Udry, S.; Benz, W., and Mordasini, C. (2008) Towards the Characterization of the Hot Neptune/Super-Earth Population around Nearby Bright Stars in: *Proceedings of the International Astronomical Union*, **4**:S253, 502–505 DOI: 10.1017/S1743921308027051 (see p. 143)

Lovis, C.; Pepe, F.; Bouchy, F.; Lo Curto, G.; Mayor, M.; Pasquini, L.; Queloz, D.; Rupprecht, G.; Udry, S., and Zucker, S. (2006) The Exoplanet Hunter HARPS: Unequalled Accuracy and Perspectives toward 1 $cm \cdot s^{-1}$ Precision in: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, **6269**: DOI: 10.1117/12.669991 (see p. 45)

Lovis, C. and Seager, S. (2010) "Radial Velocity Techniques for Exoplanets" in: *Exoplanets* (see p. 40)

Lovis, C.; Ségransan, D.; Mayor, M.; Udry, S.; Benz, W.; Bertaux, J.-L.; Bouchy, F.; Correia, A. C. M.; Laskar, J.; Lo Curto, G.; Mordasini, C.; Pepe, F.; Queloz, D., and Santos, N. C. (2011b) The HARPS

Search for Southern Extra-Solar Planets: XXVIII. Up to Seven Planets Orbiting HD 10180: Probing the Architecture of Low-Mass Planetary Systems⋆ ⋆ ⋆ in: *Astronomy & Astrophysics*, **528**: A112 DOI: 10.1051/0004-6361/201015577 (see pp. 23, 66, 69)

Lucy, L. B. and Sweeney, M. A. (1971) Spectroscopic Binaries with Circular Orbits. in: *The Astronomical Journal*, **76**: 544–556 DOI: 10.1086/111159 (see p. 56)

Luyten, W. J. (1936) A Rediscussion of the Orbits of Seventy-Seven Spectroscopic Binaries in: *The Astrophysical Journal*, **84**: 85 DOI: 10.1086/143751 (see p. 56)

Macintosh, B. A.; Graham, J. R.; Palmer, D. W.; Doyon, R.; Dunn, J.; Gavel, D. T.; Larkin, J.; Oppenheimer, B.; Saddlemyer, L.; Sivaramakrishnan, A.; Wallace, J. Kent; Bauman, B.; Erickson, D. A.; Marois, C.; Poyneer, L. A., and Soummer, R. (2008) "The Gemini Planet Imager: From Science to Design to Construction" in: ed. by N. Hubin; C. E. Max, and P. L. Wizinowich, 701518 DOI: 10.1117/12.788083 (see p. 20)

MacKay, D. J. C. (1998) "Introduction to Gaussian Processes" in: *Neural Networks and Machine Learning* ed. by C. M. Bishop NATO ASI Series Kluwer Academic Press, 133–166 (see p. 81)

– (2003) *Information Theory, Inference, and Learning Algorithms* Cambridge, UK ; New York: Cambridge University Press (see p. 60)

– (2006) *Gaussian Process Basics* (see p. 81)

Maldonado, J. and Villaver, E. (2017) Searching for Chemical Signatures of Brown Dwarf Formation in: *ArXiv e-prints*, **1702**: arXiv:1702.02904 (see p. 98)

Mallik, S. V. (1997) The Ca II Triplet Lines as Diagnostics of Luminosity, Metallicity and Chromospheric Activity in Cool Stars in: *Astronomy and Astrophysics Supplement Series*, **124**:2, 359–384 DOI: 10.1051/aas:1997199 (see p. 79)

Mamajek, E. E. and Hillenbrand, L. A. (2008) Improved Age Estimation for Solar-Type Dwarfs Using Activity-Rotation Diagnostics in: *The Astrophysical Journal*, **687**:2, 1264 (see pp. 87, 118, 123, 137, 185)

Marcy, G. W. and Butler, R. Paul (1996) A Planetary Companion to 70 Virginis in: *The Astrophysical Journal Letters*, **464**:2, L147 (see p. 18)

Marcy, G. W.; Butler, R. Paul; Fischer, D.; Vogt, S. S.; Lissauer, J. J., and Rivera, E. J. (2001) A Pair of Resonant Planets Orbiting GJ 876 in: *The Astrophysical Journal*, **556**:1, 296 DOI: 10.1086/321552 (see p. 18)

Marcy, G. W.; Butler, R. Paul; Vogt, S. S.; Fischer, D., and Lissauer, J. J. (1998) A Planetary Companion to a Nearby M4 Dwarf, Gliese 876 in: *The Astrophysical Journal Letters*, **505**:2, L147 DOI: 10.1086/311623 (see p. 18)

Marcy, G. W.; Cochran, W. D., and Mayor, M. (2000) Extrasolar Planets around Main-Sequence Stars in: *Protostars and Planets IV*, 1285 (see p. 65)

Marcy, G. W. et al. (2014) Masses, Radii, and Orbits of Small Kepler Planets: The Transition from Gaseous to Rocky Planets in: *The Astrophysical Journal Supplement Series*, **210**:2, 20 DOI: 10.1088/0067-0049/210/2/20 (see pp. 19, 25)

Marcy, G.; Butler, R. Paul; Fischer, D.; Vogt, S.; Wright, J. T.; Tinney, C. G., and Jones, H. R. A. (2005) Observed Properties of Exoplanets: Masses, Orbits, and Metallicities in: *Progress of Theoretical Physics Supplement*, **158**: 24–42 DOI: 10.1143/PTPS.158.24 (see p. 22)

Marois, C.; Macintosh, B.; Barman, T.; Zuckerman, B.; Song, I.; Patience, J.; Lafrenière, David, and Doyon, R. (2008) Direct Imaging of Multiple Planets Orbiting the Star HR 8799 in: *Science*, **322**:5906, 1348–1352 DOI: 10.1126/science.1166585 (see p. 20)

Marquardt, D. W. (1963) An Algorithm for Least-Squares Estimation of Nonlinear Parameters in: *Journal of the Society for Industrial and Applied Mathematics*, **11**:2, 431–441 DOI: 10.1137/0111030 (see p. 46)

Martin, R. G.; Livio, M., and Palaniswamy, D. (2016) Why Are Pulsar Planets Rare? in: *The Astrophysical Journal*, **832**:2, 122 DOI: 10.3847/0004-637X/832/2/122 (see p. 18)

Martínez-Arnáiz, R.; Maldonado, J.; Montes, D.; Eiroa, C., and Montesinos, B. (2010) Chromospheric Activity and Rotation of FGK Stars in the Solar Vicinity - An Estimation of the Radial Velocity Jitter in: *Astronomy & Astrophysics*, **520**: A79 DOI: 10.1051/0004-6361/200913725 (see p. 79)

Mayor, M. (1998) *The Planet around the Very Metal Rich Star 14 Herculis* Oral Presentation, Protostars & Planets IV (see p. 65)

Mayor, M.; Lovis, C., and Santos, N. C. (2014) Doppler Spectroscopy as a Path to the Detection of Earth-like Planets in: *Nature*, **513**:7518, 328–335 DOI: 10.1038/nature13780 (see p. 25)

Mayor, M.; Pepe, F.; Queloz, D.; Bouchy, F.; Rupprecht, G.; Lo Curto, G.; Avila, G.; Benz, W.; Bertaux, J.-L.; Bonfils, X.; Dall, Th; Dekker, H.; Delabre, B.; Eckert, W.; Fleury, M.; Gilliotte, A.; Gojak, D.; Guzman, J. C.; Kohler, D.; Lizon, J.-L.; Longinotti, A.; Lovis, C.; Megevand, D.; Pasquini, L.; Reyes, J.; Sivan, J.-P.; Sosnowska, D.; Soto, R.; Udry, S.; van Kesteren, A.; Weber, L., and Weilenmann, U. (2003) Setting New Standards with HARPS in: *The Messenger*, **114**: 20–24 (see pp. 40, 44, 45, 97)

Mayor, M. and Queloz, D. (1995) A Jupiter-Mass Companion to a Solar-Type Star in: *Nature*, **378**:6555, 355–359 DOI: 10.1038/378355a0 (see pp. 18, 20, 29)

Mayor, M.; Udry, S.; Naef, D.; Pepe, F.; Queloz, D.; Santos, N. C., and Burnet, M. (2004) The CORALIE Survey for Southern Extra-Solar Planets: XII. Orbital Solutions for 16 Extra-Solar Planets Discovered with CORALIE in: *Astronomy and Astrophysics*, **415**:1, 391–402 DOI: 10.1051/0004-6361:20034250 (see p. 97)

Mayor, M.; Marmier, M.; Lovis, Ch; Udry, S.; Ségransan, D.; Pepe, F.; Benz, W.; Bertaux, J.-L.; Bouchy, F.; Dumusque, X., et al. (2011) The HARPS Search for Southern Extra-Solar Planets XXXIV. Occurrence, Mass Distribution and Orbital Properties of Super-Earths and Neptune-Mass Planets in: arXiv: 1109.2497 (see pp. 22, 23, 25, 26, 45, 101, 139, 141–143, 148)

McArthur, B. E.; Endl, M.; Cochran, W. D.; Benedict, G. Fritz; Fischer, D. A.; Marcy, G. W.; Butler, R. Paul; Naef, D.; Mayor, M.; Queloz, D.; Udry, S., and Harrison, T. E. (2004) Detection of a Neptune-Mass Planet in the $P1$ Cancri System Using the Hobby-Eberly Telescope in: *The Astrophysical Journal Letters*, **614**:1, L81 DOI: 10.1086/425561 (see p. 18)

McQuillan, A.; Mazeh, T., and Aigrain, S. (2014) ROTATION PERIODS OF 34,030 *KEPLER* MAIN-SEQUENCE STARS: THE FULL AUTOCORRELATION SAMPLE in: *The Astrophysical Journal Supplement Series*, **211**:2, 24 DOI: 10.1088/0067-0049/211/2/24 (see p. 87)

Melo, C.; Santos, N. C.; Gieren, W.; Pietrzynski, G.; Ruiz, M. T.; Sousa, S. G.; Bouchy, F.; Lovis, C.; Mayor, M.; Pepe, F.; Queloz, D.; da Silva, R., and Udry, S. (2007) A New Neptune-Mass Planet Orbiting HD 219828 in: *Astronomy and Astrophysics*, **467**: 721–727 DOI: 10.1051/0004-6361:20066845 (see pp. 77, 97)

Meunier, N.; Desort, M., and Lagrange, A.-M. (2010a) Using the Sun to Estimate Earth-like Planets Detection Capabilities . II. Impact of Plages in: *Astronomy and Astrophysics*, **512**: A39 DOI: 10.1051/0004-6361/200913551 (see pp. 73, 74)

Meunier, N. and Lagrange, A.-M. (2013) Using the Sun to Estimate Earth-like Planets Detection Capabilities. IV. Correcting for the Convective Component in: *Astronomy and Astrophysics*, **551**: A101 DOI: 10.1051/0004-6361/201219917 (see p. 74)

Meunier, N.; Lagrange, A.-M., and Desort, M. (2010b) Reconstructing the Solar Integrated Radial Velocity Using MDI/SOHO in: *Astronomy and Astrophysics*, **519**: A66 DOI: 10.1051/0004-6361/201014199 (see p. 24)

Montenbruck, O. and Pfleger, T. (2000) *Astronomy on the Personal Computer* 4th Berlin ; New York: Springer (see p. 45)

Mordasini, C.; Alibert, Y., and Benz, W. (2009a) Extrasolar Planet Population Synthesis. I. Method, Formation Tracks, a. Mass-Distance Distribution in: *Astronomy and Astrophysics*, **501**:3, 1139–1160 DOI: 10.1051/0004-6361/200810301 (see p. 27)

Mordasini, C.; Alibert, Y.; Benz, W.; Klahr, H., and Henning, T. (2012) Extrasolar Planet Population Synthesis: IV. Correlations with Disk Metallicity, Mass, and Lifetime in: *Astronomy & Astrophysics*, **541**: A97 DOI: 10.1051/0004-6361/201117350 (see pp. 23, 25, 27, 148)

Mordasini, C.; Alibert, Y.; Benz, W., and Naef, D. (2009b) Extrasolar Planet Population Synthesis. II. Statistical Comparison with Observations in: *Astronomy and Astrophysics*, **501**:3, 1161–1184 DOI: 10.1051/0004-6361/200810697 (see p. 25)

Mordasini, C.; Mollière, P.; Dittkrist, K.-M.; Jin, S., and Alibert, Y. (2015) Global Models of Planet Formation and Evolution in: *International Journal of Astrobiology*, **14**:02, 201–232 DOI: 10.1017/S1473550414000263 (see p. 23)

Mortier, A.; Faria, J. P.; Correia, C. M.; Santerne, A., and Santos, N. C. (2015) BGLS: A Bayesian Formalism for the Generalised Lomb-Scargle Periodogram in: *Astronomy and Astrophysics*, **573**: A101 DOI: 10.1051/0004-6361/201424908 (see pp. 47, 162)

Mortier, A.; Faria, J. P.; Santos, N. C.; Rajpaul, V.; Figueira, P.; Boisse, I.; Collier Cameron, A.; Dumusque, X.; Lo Curto, G.; Lovis, C.; Mayor, M.; Melo, C.; Pepe, F.; Queloz, D.; Santerne, A.; Ségransan, D.; Sousa, S. G.; Sozzetti, A., and Udry, S. (2016) The HARPS Search for Southern Extra-Solar Planets: XXXIX. HD 175607, the Most Metal-Poor G Dwarf with an Orbiting Sub-Neptune in: *Astronomy & Astrophysics*, **585**: A135 DOI: 10.1051/0004-6361/201526905 (see pp. 118, 119, 121, 122, 147)

Mortier, A.; Santos, N. C.; Sousa, S.; Israelian, G.; Mayor, M., and Udry, S. (2013) On the Functional Form of the Metallicity-Giant Planet Correlation in: *Astronomy & Astrophysics*, **551**: A112 DOI: 10.1051/0004-6361/201220707 (see p. 25)

Mortier, A.; Santos, N. C.; Sozzetti, A.; Mayor, M.; Latham, D.; Bonfils, X., and Udry, S. (2012) The Frequency of Giant Planets around Metal-Poor Stars in: *Astronomy & Astrophysics*, **543**: A45 DOI: 10.1051/0004-6361/201118651 (see pp. 139, 141, 148)

Moulton, F. Ray (1970) *An Introduction to Celestial Mechanics* 2., rev. ed OCLC: 43537449 New York: Dover (see p. 29)

Moutou, C.; Mayor, M.; Lo Curto, G.; Ségransan, D.; Udry, S.; Bouchy, F.; Benz, W.; Lovis, C.; Naef, D.; Pepe, F.; Queloz, D.; Santos, N. C., and Sousa, S. G. (2011) The HARPS Search for Southern Extra-Solar Planets: XXVII. Seven New Planetary Systems in: *Astronomy & Astrophysics*, **527**: A63 DOI: 10.1051/0004-6361/201015371 (see p. 48)

Muraki, Y. et al. (2011) Discovery and Mass Measurements of a Cold, 10 Earth Mass Planet and Its Host Star in: *The Astrophysical Journal*, **741**:1, 22 DOI: 10.1088/0004-637X/741/1/22 (see p. 20)

Murison, M. A. (2006) *A Practical Method for Solving the Kepler Equation* tech. rep. U.S. Naval Observatory, Washington, DC (see p. 33)

Murray, C. D. and Correia, A. C. M. (2010) "Keplerian Orbits and Dynamics of Exoplanets" in: *Exoplanets*, 15–23 (see pp. 29–31, 34)

Naef, D.; Mayor, M.; Beuzit, J. L.; Perrier, C.; Queloz, D.; Sivan, J. P., and Udry, S. (2004) The ELODIE Survey for Northern Extra-Solar Planets: III. Three Planetary Candidates Detected with ELODIE in: *Astronomy and Astrophysics*, **414**:1, 351–359 DOI: 10.1051/0004-6361:20034091 (see p. 66)

Nardetto, N.; Mourard, D.; Kervella, P.; Mathias, Ph; Mérand, A., and Bersier, D. (2006) High Resolution Spectroscopy for Cepheids Distance Determination - I. Line Asymmetry in: *Astronomy & Astrophysics*, **453**:1, 309–319 DOI: 10.1051/0004-6361:20054333 (see p. 78)

Nayakshin, S. (2017) Dawes Review 7: The Tidal Downsizing Hypothesis of Planet Formation in: *Publications of the Astronomical Society of Australia*, **34**: DOI: 10.1017/pasa.2016.55 (see p. 26)

Nelson, B.; Ford, E. B., and Payne, M. J. (2014) RUN DMC: An Efficient, Parallel Code For Analyzing Radial Velocity Observations Using N-Body Integrations And Differential Evolution Markov Chain Monte Carlo in: *The Astrophysical Journal Supplement Series*, **210**:1, 11 DOI: 10.1088/0067-0049/210/1/11 (see p. 39)

Neves, V.; Bonfils, X.; Santos, N. C.; Delfosse, X.; Forveille, T.; Allard, F., and Udry, S. (2014) Metallicity of M Dwarfs - IV. A High-Precision [Fe/H] and Teff Technique from High-Resolution Optical Spectra for M Dwarfs in: *Astronomy & Astrophysics*, **568**: A121 DOI: 10.1051/0004-6361/201424139 (see p. 100)

Nordlund, Åke; Stein, R. F., and Asplund, M. (2009) Solar Surface Convection in: *Living Reviews in Solar Physics*, **6**:1, 2 DOI: 10.12942/lrsp-2009-2 (see p. 73)

Noyes, R. W. (1986) "Solar and Stellar Magnetic Activity" in: *Physics of the Sun* ed. by P. A. Sturrock; T. E. Holzer; D. M. Mihalas, and R. K. Ulrich Dordrecht: Springer Netherlands, 125–153 DOI: 10.1007/978-94-009-5255-3_4 (see pp. 72, 150)

Noyes, R. W.; Hartmann, L. W.; Baliunas, S. L.; Duncan, D. K., and Vaughan, A. H. (1984a) Rotation, C., and Magnetic Activity in Lower Main-Sequence Stars in: *The Astrophysical Journal*, **279**: 763–777 DOI: 10.1086/161945 (see p. 79)

Noyes, R. W.; Weiss, N. O., and Vaughan, A. H. (1984b) The Relation between Stellar Rotation Rate and Activity Cycle Periods in: *The Astrophysical Journal*, **287**: 769–773 DOI: 10.1086/162735 (see pp. 87, 109, 118, 123, 137, 185)

Oláh, K.; Kolláth, Z.; Granzer, T.; Strassmeier, K. G.; Lanza, A. F.; Järvinen, S.; Korhonen, H.; Baliunas, S. L.; Soon, W.; Messina, S., and Cutispoto, G. (2009) Multiple and Changing Cycles of Active Stars - II. Results in: *Astronomy & Astrophysics*, **501**:2, 703–713 DOI: 10.1051/0004-6361/200811304 (see p. 76)

Oshagh, M.; Boisse, I.; Boué, G.; Montalto, M.; Santos, N. C.; Bonfils, X., and Haghighipour, N. (2013) SOAP-T: A Tool to Study the Light Curve and Radial Velocity of a System with a Transiting Planet and a Rotating Spotted Star in: *Astronomy & Astrophysics*, **549**: A35 DOI: 10.1051/0004-6361/201220173 (see pp. 24, 85)

Park, B.; Jeon, Y.; Lee, C., and Han, C. (2006) Microlensing Sensitivity to Earth-Mass Planets in the Habitable Zone in: *The Astrophysical Journal*, **643**:2, 1233–1238 DOI: 10.1086/503187 (see p. 20)

Pepe, F. A. and Lovis, C. (2008) From HARPS to CODEX: Exploring the Limits of Doppler Measurements in: *Physica Scripta*, **2008**:T130, 014007 DOI: 10.1088/0031-8949/2008/T130/014007 (see p. 149)

Pepe, F.; Cameron, A. Collier; Latham, D. W.; Molinari, E.; Udry, S.; Bonomo, A. S.; Buchhave, L. A.; Charbonneau, D.; Cosentino, R.; Dressing, C. D.; Dumusque, X.; Figueira, P.; Fiorenzano, A. F. M.; Gettel, S.; Harutyunyan, A.; Haywood, R. D.; Horne, K.; Lopez-Morales, M.; Lovis, C.; Malavolta, L.; Mayor, M.; Micela, G.; Motalebi, F.; Nascimbeni, V.; Phillips, D.; Piotto, G.; Pollacco, D.; Queloz, D.; Rice, K.; Sasselov, D.; Ségransan, D.; Sozzetti, A.; Szentgyorgyi, A., and Watson, C. A. (2013) An Earth-Sized Planet with an Earth-like Density in: *Nature*, **503**:7476, 377–380 DOI: 10.1038/nature12768 (see p. 19)

Pepe, F.; Ehrenreich, D., and Meyer, M. R. (2014a) Instrumentation for the Detection and Characterization of Exoplanets in: *Nature*, **513**:7518, 358–366 DOI: 10.1038/nature13784 (see p. 40)

Pepe, F.; Mayor, M.; Galland, F.; Naef, D.; Queloz, D.; Santos, N. C.; Udry, S., and Burnet, M. (2002a) The CORALIE Survey for Southern Extra-Solar Planets VII. Two Short-Period Saturnian Companions to HD 108147 and HD 168746 in: *Astronomy and Astrophysics*, **388**: 632–638 DOI: 10.1051/0004-6361:20020433 (see p. 42)

Pepe, F.; Mayor, M.; Rupprecht, G.; Avila, G.; Ballester, P.; Beckers, J.-L.; Benz, W.; Bertaux, J.-L.; Bouchy, F.; Buzzoni, B.; Cavadore, C.; Deiries, S.; Dekker, H.; Delabre, B.; D'Odorico, S.; Eckert, W.; Fischer, J.; Fleury, M.; George, M.; Gilliotte, A.; Gojak, D.; Guzman, J.-C.; Koch, F.; Kohler, D.; Kotzlowski, H.; Lacroix, D.; Le Merrer, J.; Lizon, J.-L.; Lo Curto, G.; Longinotti, A.; Megevand, D.; Pasquini, L.; Petitpas, P.; Pichard, M.; Queloz, D.; Reyes, J.; Richaud, P.; Sivan, J.-P.; Sosnowska, D.; Soto, R.; Udry, S.; Ureta, E.; van Kesteren, A.; Weber, L.; Weilenmann, U.; Wicenec, A.; Wieland, G.; Christensen-Dalsgaard, J.; Dravins, D.; Hatzes, A.; Kürster, M.; Paresce, F., and Penny, A. (2002b) HARPS: ESO's Coming Planet Searcher. Chasing Exoplanets with the La Silla 3.6-m Telescope in: *The Messenger*, **110**: 9–14 (see p. 44)

Pepe, F.; Molaro, P.; Cristiani, S.; Rebolo, R.; Santos, N. C.; Dekker, H.; Mégevand, D.; Zerbi, F. M.; Cabral, A.; Di Marcantonio, P., et al. (2014b) ESPRESSO: The next European Exoplanet Hunter in: *Astronomische Nachrichten*, **335**:1, 8–20 (see pp. 149, 150)

Pérez, F. and Granger, B. E. (2007) IPython: a System for Interactive Scientific Computing in: *Computing in Science and Engineering*, **9**:3, 21–29 DOI: 10.1109/MCSE.2007.53 (see p. 161)

Perrakis, K.; Ntzoufras, I., and Tsionas, E. G. (2014) On the Use of Marginal Posteriors in Marginal Likelihood Estimation via Importance Sampling in: *Computational Statistics & Data Analysis*, **77**: 54–69 DOI: 10.1016/j.csda.2014.03.004 (see p. 151)

Perryman, M. a. C.; Boer, K. S. de; Gilmore, G.; Høg, E.; Lattanzi, M. G.; Lindegren, L.; Luri, X.; Mignard, F.; Pace, O., and Zeeuw, P. T. de (2001) GAIA: Composition, Formation and Evolution of the Galaxy in: *Astronomy & Astrophysics*, **369**:1, 339–363 DOI: 10.1051/0004-6361:20010085 (see p. 21)

Perryman, M.; Hartman, J.; Bakos, Gáspár Á, and Lindegren, L. (2014) Astrometric Exoplanet Detection with Gaia in: *The Astrophysical Journal*, **797**:1, 14 DOI: 10.1088/0004-637X/797/1/14 (see p. 21)

Phillips, D. F.; Glenday, A. G.; Dumusque, X.; Buchschacher, N.; Cameron, A. Collier; Cecconi, M.; Charbonneau, D.; Cosentino, R.; Ghedina, A.; Haywood, R.; Latham, D. W.; Li, Chih-Hao; Lodi, M.; Lovis, C.; Molinari, E.; Pepe, F.; Sasselov, D.; Szentgyorgyi, A.; Udry, S., and Walsworth, R. L. (2016) "An Astro-Comb Calibrated Solar Telescope to Search for the Radial Velocity Signature of Venus" in: vol. 9912, 99126Z DOI: 10.1117/12.2232452 (see p. 150)

Pollacco, D.; Skillen, I.; Cameron, A.; Christian, D.; Irwin, J.; Lister, T.; Street, R.; West, R.; Clarkson, W.; Evans, N.; Fitzsimmons, A.; Haswell, C.; Hellier, C.; Hodgkin, S.; Horne, K.; Jones, B.; Kane, S.; Keenan, F.; Norton, A.; Osborne, J.; Ryans, R., and Wheatley, P. (2006) The WASP Project and SuperWASP Camera in: *Astrophysics and Space Science*, **304**:1-4, 253–255 DOI: 10.1007/s10509-006-9124-x (see p. 19)

Pollack, J. B.; Hubickyj, O.; Bodenheimer, P.; Lissauer, J. J.; Podolak, M., and Greenzweig, Y. (1996) Formation of the Giant Planets by Concurrent Accretion of Solids and Gas in: *Icarus*, **124**:1, 62–85 DOI: 10.1006/icar.1996.0190 (see p. 25)

Pont, F.; Aigrain, S., and Zucker, S. (2011) Reassessing the Radial-Velocity Evidence for Planets around CoRoT-7 in: *Monthly Notices of the Royal Astronomical Society*, **411**:3, 1953–1962 DOI: 10.1111/j.1365-2966.2010.17823.x (see p. 89)

Prato, L.; Huerta, M.; Johns-Krull, C. M.; Mahmud, N.; Jaffe, D. T., and Hartigan, P. (2008) A Young-Planet Search in Visible and Infrared Light: DN Tauri, V836 Tauri, and V827 Tauri in: *The Astrophysical Journal*, **687**: DOI: 10.1086/593201 (see pp. 77, 125)

Quanz, S. P.; Crossfield, I.; Meyer, M. R.; Schmalzl, E., and Held, J. (2015) Direct Detection of Exoplanets in the 3–10 *M*m Range with E-ELT/METIS in: *International Journal of Astrobiology*, **14**:02, 279–289 DOI: 10.1017/S1473550414000135 (see p. 150)

Queloz, D.; Bouchy, F.; Moutou, C.; Hatzes, A.; Hébrard, G.; Alonso, R.; Auvergne, M.; Baglin, A.; Barbieri, M.; Barge, P.; Benz, W.; Bordé, P.; Deeg, H. J.; Deleuil, M.; Dvorak, R.; Erikson, A.; Ferraz Mello, S.; Fridlund, M.; Gandolfi, D.; Gillon, M.; Guenther, E.; Guillot, T.; Jorda, L.; Hartmann, M.; Lammer, H.; Léger, A.; Llebaria, A.; Lovis, C.; Magain, P.; Mayor, M.; Mazeh, T.; Ollivier, M.; Pätzold, M.; Pepe, F.; Rauer, H.; Rouan, D.; Schneider, J.; Segransan, D.; Udry, S., and Wuchterl, G. (2009) The CoRoT-7 Planetary System: Two Orbiting Super-Earths in: *Astronomy and Astrophysics*, **506**:1, 303–319 DOI: 10.1051/0004-6361/200913096 (see pp. 19, 78, 89)

Queloz, D.; Henry, G. W.; Sivan, J. P.; Baliunas, S. L.; Beuzit, J. L.; Donahue, R. A.; Mayor, M.; Naef, D.; Perrier, C., and Udry, S. (2001) No Planet for HD 166435 in: *Astronomy and Astrophysics*, **379**:1, 279–287 DOI: 10.1051/0004-6361:20011308 (see pp. 76, 77)

Rajpaul, V.; Aigrain, S.; Osborne, M. A.; Reece, S., and Roberts, S. (2015) A Gaussian Process Framework for Modelling Stellar Activity Signals in Radial Velocity Data in: *Monthly Notices of the Royal Astronomical Society*, **452**:3, 2269–2291 DOI: 10.1093/mnras/stv1428 (see pp. 78, 118)

Rajpaul, V.; Aigrain, S., and Roberts, S. (2016) Ghost in the Time Series: No Planet for Alpha Cen B in: *Monthly Notices of the Royal Astronomical Society: Letters*, **456**:1, L6–L10 DOI: 10.1093/mnrasl/slv164 (see p. 26)

Ramírez, I.; Asplund, M.; Baumann, P.; Meléndez, J., and Bensby, T. (2010) A Possible Signature of Terrestrial Planet Formation in the Chemical Composition of Solar Analogs in: *Astronomy and Astrophysics*, **521**: A33 DOI: 10.1051/0004-6361/201014456 (see p. 27)

Rasmussen, C. Edward and Williams, C. K. I. (2006) *Gaussian Processes for Machine Learning* Cambridge: MIT Press (see pp. 80, 81, 84)

Rauer, H. et al. (2014) The PLATO 2.0 Mission in: *Experimental Astronomy*, **38**:1-2, 249–330 DOI: 10.1007/s10686-014-9383-4 (see p. 150)

Reddy, B. E.; Lambert, D. L.; Laws, C.; Gonzalez, G., and Covey, K. (2002) A Search for 6Li in Stars with Planets in: *Monthly Notices of the Royal Astronomical Society*, **335**:4, 1005–1016 DOI: 10.1046/j.1365-8711.2002.05682.x (see p. 27)

Reiners, A.; Lemke, U.; Bauer, F.; Beeck, B., and Huke, P. (2016) Radial Velocity Observations of the 2015 Mar. 20 Eclipse - A Benchmark Rossiter-McLaughlin Curve with Zero Free Parameters in: *Astronomy & Astrophysics*, **595**: A26 DOI: 10.1051/0004-6361/201629088 (see p. 150)

Reuyl, D. and Holmberg, E. (1943) On the Existence of a Third Component in the System 70 Ophiuchi. in: *The Astrophysical Journal*, **97**: 41 DOI: 10.1086/144489 (see p. 21)

Ribas, I.; Bolmont, E.; Selsis, F.; Reiners, A.; Leconte, J.; Raymond, S. N.; Engle, S. G.; Guinan, E. F.; Morin, J.; Turbet, M.; Forget, F., and Anglada-Escudé, G. (2016) The Habitability of Proxima Centauri b - I. Irradiation, Rotation and Volatile Inventory from Formation to the Present in: *Astronomy & Astrophysics*, **596**: A111 DOI: 10.1051/0004-6361/201629576 (see p. 18)

Rivera, E. J.; Laughlin, G.; Butler, R. Paul; Vogt, S. S.; Haghighipour, N., and Stefano Meschiari (2010) The Lick-Carnegie Exoplanet Survey: A Uranus-Mass Fourth Planet for GJ 876 in an Extrasolar Laplace Configuration in: *The Astrophysical Journal*, **719**:1, 890 DOI: 10.1088/0004-637X/719/1/890 (see p. 38)

Rivera, E. J.; Lissauer, J. J.; Butler, R. Paul; Marcy, G. W.; Vogt, S. S.; Fischer, D. A.; Brown, T. M.; Laughlin, G., and Henry, G. W. (2005) A 7.5 M⊕ Planet Orbiting the Nearby Star, GJ 876 in: *The Astrophysical Journal*, **634**:1, 625 DOI: 10.1086/491669 (see p. 18)

Robert, C. P.; Chopin, N., and Rousseau, J. (2009) Harold Jeffreys's Theory of Probability Revisited in: *Statistical Science*, **24**:2, 141–172 DOI: 10.1214/09-STS284 (see p. 51)

Robert P., Christian (2007) *The Bayesian Choice* Springer Texts in Statistics New York, NY: Springer (see p. 51)

Robertson, P.; Endl, M.; Cochran, W. D., and Dodson-Robinson, S. E. (2013) Hα Activity of Old M Dwarfs: Stellar Cycles and Mean Activity Levels for 93 Low-Mass Stars in the Solar Neighborhood in: *The Astrophysical Journal*, **764**: 3 DOI: 10.1088/0004-637X/764/1/3 (see pp. 79, 137, 150)

Robertson, P. and Mahadevan, S. (2014) Disentangling Planets and Stellar Activity for Gliese 667C in: *The Astrophysical Journal Letters*, **793**:2, L24 (see pp. 24, 26, 79)

Robertson, P.; Roy, A., and Mahadevan, S. (2015) Stellar Activity Mimics a Habitable-Zone Planet around Kapteyn's Star in: *The Astrophysical Journal*, **805**:2, L22 DOI: 10.1088/2041-8205/805/2/L22 (see p. 79)

Route, M. (2016) The Discovery of Solar-like Activity Cycles Beyond the End of the Main Sequence? in: *The Astrophysical Journal*, **830**: DOI: 10.3847/2041-8205/830/2/L27 (see p. 72)

Rupprecht, G.; Pepe, F.; Mayor, M.; Queloz, D.; Bouchy, F.; Avila, G.; Benz, W.; Bertaux, J.; Bonfils, X.; Dall, Th.; Delabre, B.; Dekker, H.; Eckert, W.; Fleury, M.; Gilliotte, A.; Gojak, D.; Guzman, J. C.; Kohler, D.; Lizon, J.; Lo Curto, G.; Longinotti, A.; Lovis, C.; Megevand, D.; Pasquini, L.; Reyes, J.; Sivan, J.; Sosnowska, D.; Soto, R.; Udry, S.; Van Kesteren, A.; Weber, L., and Weilenmann, U. (2004) "The Exoplanet Hunter HARPS: Performance and First Results" in: 148 DOI: 10.1117/12.551267 (see p. 45)

Rutten, R. G. M. (1984) Magnetic Structure in Cool Stars. VII - Absolute Surface Flux in CA II H and K Line Cores in: *Astronomy and Astrophysics*, **130**: 353–360 (see p. 79)

Saar, S. H.; Butler, R. Paul, and Marcy, G. W. (1998) Magnetic Activity-Related Radial Velocity Variations in Cool Stars: First Results from the Lick Extrasolar Planet Survey in: *The Astrophysical Journal Letters*, **498**: L153–L157 DOI: 10.1086/311325 (see p. 79)

Saar, S. H. and Donahue, R. A. (1997) Activity-Related Radial Velocity Variation in Cool Stars in: *The Astrophysical Journal*, **485**:1, 319 (see p. 73)

Santerne, A.; Beaulieu, J.-P.; Ayala, B. Rojas; Boisse, I.; Schlawin, E.; Almenara, J.-M.; Batista, V.; Bennett, D.; Díaz, R. F.; Figueira, P.; James, D. J.; Herter, T.; Lillo-Box, J.; Marquette, J. B.; Ranc, C.; Santos, N. C., and Sousa, S. G. (2016a) Spectroscopic Characterisation of Microlensing Events - Towards a New Interpretation of OGLE-2011-BLG-0417 in: *Astronomy & Astrophysics*, **595**: L11 DOI: 10.1051/0004-6361/201527710 (see p. 20)

Santerne, A.; Díaz, R. F.; Almenara, J.-M.; Bouchy, F.; Deleuil, M.; Figueira, P.; Hébrard, G.; Moutou, C.; Rodionov, S., and Santos, N. C. (2015) PASTIS: Bayesian Extrasolar Planet Validation – II. Constraining Exoplanet Blend Scenarios Using Spectroscopic Diagnoses in: *Monthly Notices of the Royal Astronomical Society*, **451**:3, 2337–2351 DOI: 10.1093/mnras/stv1080 (see p. 78)

Santerne, A.; Fressin, F.; Díaz, R. F.; Figueira, P.; Almenara, J.-M., and Santos, N. C. (2013) The Contribution of Secondary Eclipses as Astrophysical False Positives to Exoplanet Transit Surveys in: *Astronomy & Astrophysics*, **557**: A139 DOI: 10.1051/0004-6361/201321475 (see p. 19)

Santerne, A.; Moutou, C.; Tsantaki, M.; Bouchy, F.; Hébrard, G.; Adibekyan, V.; Almenara, J.-M.; Amard, L.; Barros, S. C. C.; Boisse, I.; Bonomo, A. S.; Bruno, G.; Courcol, B.; Deleuil, M.; Demangeon, O.; Díaz, R. F.; Guillot, T.; Havel, M.; Montagnier, G.; Rajpurohit, A. S.; Rey, J., and Santos, N. C. (2016b) SOPHIE Velocimetry of Kepler Transit Candidates - XVII. The Physical Properties of Giant Exoplanets within 400 Days of Period in: *Astronomy & Astrophysics*, **587**: A64 DOI: 10.1051/0004-6361/201527329 (see p. 57)

Santos, N. C. (2002) *New Studies on Stars with Giant Planets*. PhD thesis Université de Genève (see p. 100)

Santos, N. C.; Adibekyan, V.; Figueira, P.; Andreasen, D. T.; Barros, S. C. C.; Delgado-Mena, E.; Demangeon, O.; Faria, J. P.; Oshagh, M.; Sousa, S. G.; Viana, P. T. P., and Ferreira, A. C. S. (2017) Observational Evidence for Two Distinct Giant Planet Populations in: *Astronomy & Astrophysics*, **603**: A30 DOI: 10.1051/0004-6361/201730761 arXiv: 1705.06090 (see pp. 26, 98)

Santos, N. C.; Adibekyan, V.; Mordasini, C.; Benz, W.; Delgado-Mena, E.; Dorn, C.; Buchhave, L.; Figueira, P.; Mortier, A.; Pepe, F.; Santerne, A.; Sousa, S. G., and Udry, S. (2015) Constraining Planet Structure from Stellar Chemistry: The Cases of CoRoT-7, Kepler-10, and Kepler-93 in: *Astronomy & Astrophysics*, **580**: L13 DOI: 10.1051/0004-6361/201526850 (see pp. 25, 27)

Santos, N. C.; Bouchy, F.; Mayor, M.; Pepe, F.; Queloz, D.; Udry, S.; Lovis, C.; Bazot, M.; Benz, W.; Bertaux, J.-L.; Curto, G. Lo; Delfosse, X.; Mordasini, C.; Naef, D.; Sivan, J.-P., and Vauclair, S. (2004a) The HARPS Survey for Southern Extra-Solar Planets - II. A 14 Earth-Masses Exoplanet around μ Arae in: *Astronomy & Astrophysics*, **426**:1, L19 DOI: 10.1051/0004-6361:200400076 (see p. 18)

Santos, N. C.; Gomes da Silva, J.; Lovis, C., and Melo, C. (2010a) Do Stellar Magnetic Cycles Influence the Measurement of Precise Radial Velocities? in: *Astronomy and Astrophysics*, **511**: A54 DOI: 10.1051/0004-6361/200913433 (see pp. 76, 80)

Santos, N. C.; Israelian, G., and Mayor, M. (2001) The Metal-Rich Nature of Stars with Planets in: *Astronomy and Astrophysics*, **373**:3, 1019–1031 DOI: 10.1051/0004-6361:20010648 (see p. 25)

– (2004b) Spectroscopic [Fe/H] for 98 Extra-Solar Planet-Host Stars: Exploring the Probability of Planet Formation in: *Astronomy and Astrophysics*, **415**:3, 1153–1166 DOI: 10.1051/0004-6361:20034469 (see pp. 25, 100, 148)

Santos, N. C.; Mayor, M.; Benz, W.; Bouchy, F.; Figueira, P.; Lo Curto, G.; Lovis, C.; Melo, C.; Moutou, C.; Naef, D.; Pepe, F.; Queloz, D.; Sousa, S. G., and Udry, S. (2010b) The HARPS Search for Southern Extra-Solar Planets: XXI. Three New Giant Planets Orbiting the Metal-Poor Stars HD 5388, HD 181720, and HD 190984 in: *Astronomy and Astrophysics*, **512**: A47 DOI: 10.1051/0004-6361/200913489 (see pp. 24, 97, 108, 109, 112, 113, 116)

Santos, N. C.; Mayor, M.; Bonfils, X.; Dumusque, X.; Bouchy, F.; Figueira, P.; Lovis, C.; Melo, C.; Pepe, F.; Queloz, D.; Ségransan, D.; Sousa, S. G., and Udry, S. (2011) The HARPS Search for Southern Extrasolar Planets: XXV. Results from the Metal-Poor Sample in: *Astronomy & Astrophysics*, **526**: A112 DOI: 10.1051/0004-6361/201015494 (see pp. 98, 108, 110, 113)

Santos, N. C.; Mayor, M.; Bouchy, F.; Pepe, F.; Queloz, D., and Udry, S. (2007) The HARPS Search for Southern Extra-Solar Planets. XII. A Giant Planet Orbiting the Metal-Poor Star HD 171028 in: *Astronomy and Astrophysics*, **474**:2, 647–651 DOI: 10.1051/0004-6361:20078129 (see pp. 97, 108, 110, 116)

Santos, N. C.; Mayor, M.; Naef, D.; Pepe, F.; Queloz, D.; Udry, S., and Blecha, A. (2000) The CORALIE Survey for Southern Extra-Solar Planets. IV. Intrinsic Stellar Limitations to Planet Searches with Radial-Velocity Techniques in: *Astronomy and Astrophysics*, **361**: 265–272 (see p. 79)

Santos, N. C.; Mayor, M.; Naef, D.; Pepe, F.; Queloz, D.; Udry, S.; Burnet, M.; Clausen, J. V.; Helt, B. E.; Olsen, E. H., and Pritchard, J. D. (2002) The CORALIE Survey for Southern Extra-Solar Planets. IX. A 1.3-Day Period Brown Dwarf Disguised as a Planet in: *Astronomy and Astrophysics*, **392**: 215–229 DOI: 10.1051/0004-6361:20020876 (see pp. 77, 98, 101)

Santos, N. C.; Mortier, A.; Faria, J. P.; Dumusque, X.; Adibekyan, V. Zh.; Delgado-Mena, E.; Figueira, P.; Benamati, L.; Boisse, I.; Cunha, D.; Gomes da Silva, J.; Lo Curto, G.; Lovis, C.; Martins, J. H. C.; Mayor, M.; Melo, C.; Oshagh, M.; Pepe, F.; Queloz, D.; Santerne, A.; Ségransan, D.; Sozzetti, A.; Sousa, S. G., and Udry, S. (2014) The HARPS Search for Southern Extra-Solar Planets: XXXV. The Interesting Case of HD 41248: Stellar Activity, No Planets? in: *Astronomy & Astrophysics*, **566**: A35 DOI: 10.1051/0004-6361/201423808 (see pp. 24, 78, 98, 122–124, 147)

Santos, N. C.; Santerne, A.; Faria, J. P.; Rey, J.; Correia, A. C. M.; Laskar, J.; Udry, S.; Adibekyan, V.; Bouchy, F.; Delgado-Mena, E.; Melo, C.; Dumusque, X.; Hébrard, G.; Lovis, C.; Mayor, M.; Montalto, M.; Mortier, A.; Pepe, F.; Figueira, P.; Sahlmann, J.; Ségransan, D., and Sousa, S. G. (2016) An Extreme Planetary System around HD 219828: One Long-Period Super Jupiter to a Hot-Neptune Host Star⋆ ⋆ ⋆ in: *Astronomy & Astrophysics*, **592**: A13 DOI: 10.1051/0004-6361/201628374 (see p. 21)

Santos, N. C.; Sousa, S. G.; Mortier, A.; Neves, V.; Adibekyan, V.; Tsantaki, M.; Mena, E. Delgado; Bonfils, X.; Israelian, G.; Mayor, M., and Udry, S. (2013) SWEET-Cat: A Catalogue of Parameters for Stars With ExoplanETs - I. New Atmospheric Parameters and Masses for 48 Stars with Planets in: *Astronomy & Astrophysics*, **556**: A150 DOI: 10.1051/0004-6361/201321286 (see p. 101)

Santos, N. C.; Udry, S.; Mayor, M.; Naef, D.; Pepe, F.; Queloz, D.; Burki, G.; Cramer, N., and Nicolet, B. (2003) The CORALIE Survey for Southern Extra-Solar Planets. XI. The Return of the Giant Planet Orbiting HD 192263 in: *Astronomy and Astrophysics*, **406**: 373–381 DOI: 10.1051/0004-6361:20030776 (see p. 77)

Savage, L. J. (1972) *The Foundations of Statistics* 2d rev. ed New York: Dover Publications (see p. 51)

Scargle, J. D. (1982) Studies in Astronomical Time Series Analysis. II - Statistical Aspects of Spectral Analysis of Unevenly Spaced Data in: *The Astrophysical Journal*, **263**: 835 DOI: 10.1086/160554 (see pp. 47, 162)

Schneider, J.; Dedieu, C.; Le Sidaner, P.; Savalle, R., and Zolotukhin, I. (2011) Defining and Cataloging Exoplanets: The Exoplanet.Eu Database in: *Astronomy and Astrophysics*, **532**: A79 DOI: 10.1051/0004-6361/201116713 (see p. 17)

Schrijver, C. J and Zwaan, C (2008) *Solar and Stellar Magnetic Activity* OCLC: 913071749 Cambridge (Reino Unido); New York: Cambridge University Press (see p. 72)

Schrijver, C.j. (2002) Solar Spots as Prototypes for Stellar Spots in: *Astronomische Nachrichten*, **323**:3-4, 157–164 DOI: 10.1002/1521-3994(200208)323:3/4<157::AID-ASNA157>3.0.CO;2-Q (see pp. 74, 87, 150)

Schwarz, G. (1978) Estimating the Dimension of a Model in: *The Annals of Statistics*, **6**:2, 461–464 DOI: 10.1214/aos/1176344136 (see p. 49)

Setiawan, J.; Hatzes, A. P.; Lühe, O. von der; Pasquini, L.; Naef, D.; Silva, L. da; Udry, S.; Queloz, D., and Girardi, L. (2003) Evidence of a Sub-Stellar Companion around HD 47536 in: *Astronomy & Astrophysics*, **398**:2, L19–L23 DOI: 10.1051/0004-6361:20021846 (see p. 97)

Setiawan, J.; Weise, P.; Henning, Th.; Hatzes, A. P.; Pasquini, L.; da Silva, L.; Girardi, L.; von der Lühe, O.; Döllinger, M. P.; Weiss, A., and Biazzo, K. (2008) "Planets Around Active Stars" in: *Precision Spectroscopy in Astrophysics* ed. by B. Leibundgut; N. C. Santos; L. Pasquini; A. C. M. Correia, and M. Romaniello Berlin, H.: Springer Berlin Heidelberg, 201–204 (see p. 97)

Shannon, C. Elwood and Weaver, W. (1998) *The Mathematical Theory of Communication* OCLC: 246600266 Urbana: Univ. of Illinois Press (see p. 60)

Shapiro, A. I.; Schmutz, W.; Cessateur, G., and Rozanov, E. (2013) The Place of the Sun among the Sun-like Stars in: *Astronomy & Astrophysics*, **552**: A114 DOI: 10.1051/0004-6361/201220512 (see p. 80)

Shapiro, A. I.; Solanki, S. K.; Krivova, N. A.; Yeo, K. L., and Schmutz, W. K. (2016) Are Solar Brightness Variations Faculae- or Spot-Dominated? in: *Astronomy & Astrophysics*, **589**: A46 DOI: 10.1051/0004-6361/201527527 (see p. 74)

Shen, Y. and Turner, E. L. (2008) On the Eccentricity Distribution of Exoplanets from Radial Velocity Surveys in: *The Astrophysical Journal*, **685**:1, 553 DOI: 10.1086/590548 (see p. 57)

Sivia, D. S and Skilling, J (2006) *Data Analysis: A Bayesian Tutorial* Oxford: Oxford University Press (see p. 154)

Skilling, J. (2004) "Nested Sampling" in: *Bayesian Inference and Maximum Entropy Methods in Science and Engineering* vol. 735 Garching (Germany), 395–405 DOI: 10.1063/1.1835238 (see p. 151)

– (2006) Nested Sampling for General Bayesian Computation in: *Bayesian Analysis*, **1**:4, 833–859 DOI: 10.1214/06-BA127 (see pp. 151, 152)

Sneden, C.; Bean, J.; Ivans, I.; Lucatello, S., and Sobeck, J. (2012) MOOG: LTE Line Analysis and Spectrum Synthesis in: *Astrophysics Source Code Library*, ascl:1202.009 (see p. 100)

Solanki, S. K. (2003) Sunspots: An Overview in: *The Astronomy and Astrophysics Review*, **11**:2-3, 153–286 DOI: 10.1007/s00159-003-0018-4 (see p. 73)

Soto, M. G.; Jenkins, J. S., and Jones, M. I. (2015) RAFT – I. Discovery of New Planetary Candidates and Updated Orbits from Archival FEROS Spectra in: *Monthly Notices of the Royal Astronomical Society*, **451**:3, 3131–3144 DOI: 10.1093/mnras/stv1144 (see p. 97)

Sousa, S. G. (2014) "ARES + MOOG: A Practical Overview of an Equivalent Width (EW) Method to Derive Stellar Parameters" in: *Determination of Atmospheric Parameters of B-, A-, F- and G-Type Stars* ed. by E. Niemczura; B. Smalley, and W. Pych Cham: Springer International Publishing, 297–310 DOI: 10.1007/978-3-319-06956-2_26 (see p. 100)

Sousa, S. G.; Santos, N. C.; Israelian, G.; Lovis, C.; Mayor, M.; Silva, P. B., and Udry, S. (2011a) Spectroscopic Characterization of a Sample of Metal-Poor Solar-Type Stars from the HARPS Planet Search Program: Precise Spectroscopic Parameters and Mass Estimation⋆⋆ in: *Astronomy & Astrophysics*, **526**: A99 DOI: 10.1051/0004-6361/201015646 (see pp. 100, 122)

Sousa, S. G.; Santos, N. C.; Israelian, G.; Mayor, M., and Monteiro, M. J. P. F. G. (2006) Spectroscopic Parameters for a Sample of Metal-Rich Solar-Type Stars in: *Astronomy & Astrophysics*, **458**:3, 873–880 DOI: 10.1051/0004-6361:20065658 (see p. 100)

– (2007) A New Code for Automatic Determination of Equivalent Widths: Automatic Routine for Line Equivalent Widths in Stellar Spectra (ARES) in: *Astronomy & Astrophysics*, **469**:2, 783–791 DOI: 10.1051/0004-6361:20077288 (see p. 100)

Sousa, S. G.; Santos, N. C.; Israelian, G.; Mayor, M., and Udry, S. (2011b) Spectroscopic Stellar Parameters for 582 FGK Stars in the HARPS Volume-Limited Sample: Revising the Metallicity-Planet Correlation⋆ in: *Astronomy & Astrophysics*, **533**: A141 DOI: 10.1051/0004-6361/201117699 (see pp. 25, 26, 101, 148)

Sousa, S. G.; Santos, N. C.; Mayor, M.; Udry, S.; Casagrande, L.; Israelian, G.; Pepe, F.; Queloz, D., and Monteiro, M. J. P. F. G. (2008) Spectroscopic Parameters for 451 Stars in the HARPS GTO Planet Search Program in: *Astronomy and Astrophysics*, **487**:1, 373–381 DOI: 10.1051/0004-6361:200809698 (see p. 100)

Sozzetti, A.; Casertano, S.; Lattanzi, M. G., and Spagna, A. (2001) Detection and Measurement of Planetary Systems with GAIA in: *Astronomy & Astrophysics*, **373**:3, L21–L24 DOI: 10.1051/0004-6361:20010788 (see p. 21)

Sozzetti, A.; Torres, G.; Latham, D. W.; Carney, B. W.; Stefanik, R. P.; Boss, A. P.; Laird, J. B., and Korzennik, S. G. (2006) A Keck HIRES Doppler Search for Planets Orbiting Metal-Poor Dwarfs. I. Testing Giant Planet Formation and Migration Scenarios in: *The Astrophysical Journal*, **649**:1, 428 DOI: 10.1086/506267 (see p. 97)

Sozzetti, A.; Torres, G.; Latham, D. W.; Stefanik, R. P.; Korzennik, S. G.; Boss, A. P.; Carney, B. W., and Laird, J. B. (2009) A Keck HIRES Doppler Search for Planets Orbiting Metal-Poor Dwarfs. II. On the Frequency of Giant Planets in the Metal-Poor Regime in: *The Astrophysical Journal*, **697**: 544–556 DOI: 10.1088/0004-637X/697/1/544 (see p. 139)

Spiegel, D. S.; Burrows, A., and Milsom, J. A. (2011) The Deuterium-Burning Mass Limit for Brown Dwarfs and Giant Planets in: *The Astrophysical Journal*, **727**:1, 57 DOI: 10.1088/0004-637X/727/1/57 (see pp. 17, 106)

Spiegel, E. A. and Zahn, J.-P. (1992) The Solar Tachocline in: *Astronomy and Astrophysics*, **265**: 106–114 (see p. 71)

Stephens, M. (2000) Bayesian Analysis of Mixture Models with an Unknown Number of Components—an Alternative to Reversible Jump Methods in: *The Annals of Statistics*, **28**:1, 40–74 DOI: 10.1214/aos/1016120364 (see pp. 53, 156)

Stepinski, T. F.; Malhotra, R., and Black, D. C. (2000) The *υ* Andromedae System: Models and Stability in: *The Astrophysical Journal*, **545**:2, 1044 DOI: 10.1086/317829 (see p. 47)

Stoica, P. and Moses, R. L. (2005) *Spectral Analysis of Signals* Upper Saddle River, N.J: Pearson/Prentice Hall (see p. 72)

Strand, K. Aa. (1943) 61 Cygni as a Triple System in: *Publications of the Astronomical Society of the Pacific*, **55**: 29 DOI: 10.1086/125484 (see p. 21)

Struve, O. (1952) Proposal for a Project of High-Precision Stellar Radial Velocity Work in: *The Observatory*, **72**: 199–200 (see p. 29)

Sumi, T.; Kamiya, K.; Bennett, D. P.; Bond, I. A.; Abe, F.; Botzler, C. S.; Fukui, A.; Furusawa, K.; Hearnshaw, J. B.; Itow, Y.; Kilmartin, P. M.; Korpela, A.; Lin, W.; Ling, C. H.; Masuda, K.; Matsubara, Y.; Miyake, N.; Motomura, M.; Muraki, Y.; Nagaya, M.; Nakamura, S.; Ohnishi, K.; Okumura, T.; Perrott, Y. C.; Rattenbury, N.; Saito, To; Sako, T.; Sullivan, D. J.; Sweatman, W. L.; Tristram, P. J.; Udalski, A.; Szymański, M. K.; Kubiak, M.; Pietrzyński, G.; Poleski, R.; Soszyński, I.; Wyrzykowski, Ł; Ulaczyk, K., and MOA Collaboration (2011) Unbound or Distant Planetary Mass Population Detected by Gravitational Microlensing in: *Nature*, **473**:7347, 349–352 DOI: 10.1038/nature10092 (see p. 20)

Suzuki, D.; Bennett, D. P.; Sumi, T.; Bond, I. A.; Rogers, L. A.; Abe, F.; Asakura, Y.; Bhattacharya, A.; Donachie, M.; Freeman, M.; Fukui, A.; Hirao, Y.; Itow, Y.; Koshimoto, N.; Li, M. C. A.; Ling, C. H.; Masuda, K.; Matsubara, Y.; Muraki, Y.; Nagakane, M.; Onishi, K.; Oyokawa, H.; Rattenbury, N.; Saito, To.; Sharan, A.; Shibai, H.; Sullivan, D. J.; Tristram, P. J.; Yonehara, A., and MOA Collaboration (2016) The Exoplanet Mass-Ratio Function from the MOA-II Survey: Discovery of a Break and Likely Peak at a Neptune Mass in: *The Astrophysical Journal*, **833**: 145 DOI: 10.3847/1538-4357/833/2/145 (see p. 20)

Thomas, J. H and Weiss, N. O (2012) *Sunspots and Starspots* OCLC: 797977801 Cambridge: Cambridge University Press (see pp. 73, 74)

Tinney, C. G.; Butler, R. Paul; Marcy, G. W.; Jones, H. R. A.; Penny, A. J.; McCarthy, C., and Carter, B. D. (2002) Two Extrasolar Planets from the Anglo-Australian Planet Search in: *The Astrophysical Journal*, **571**:1, 528–531 DOI: 10.1086/339916 (see p. 97)

Title, A. M.; Tarbell, T. D., and Topka, K. P. (1987) On the Relation between Magnetic Field Structures and Granulation in: *The Astrophysical Journal*, **317**: 892–899 DOI: 10.1086/165339 (see p. 74)

Tobias, S. M. (2002) The Solar Dynamo in: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **360**:1801, 2741–2756 DOI: 10.1098/rsta.2002.1090 (see p. 71)

Torres, G.; Andersen, J., and Giménez, A. (2010) Accurate Masses and Radii of Normal Stars: Modern Results and Applications in: *Astronomy and Astrophysics Review*, **18**: 67–126 DOI: 10.1007/s00159-009-0025-1 (see p. 101)

Torres, G.; Konacki, M.; Sasselov, D. D., and Jha, S. (2004) Testing Blend Scenarios for Extrasolar Transiting Planet Candidates. I. OGLE-TR-33: A False Positive in: *The Astrophysical Journal*, **614**:2, 979 DOI: 10.1086/423734 (see p. 19)

Torres, G.; Winn, J. N., and Holman, M. J. (2008) Improved Parameters for Extrasolar Transiting Planets in: *The Astrophysical Journal*, **677**:2, 1324–1342 DOI: 10.1086/529429 (see pp. 25, 37)

Tsantaki, M.; Sousa, S. G.; Adibekyan, V. Zh; Santos, N. C.; Mortier, A., and Israelian, G. (2013) Deriving Precise Parameters for Cool Solar-Type Stars - Optimizing the Iron Line List in: *Astronomy & Astrophysics*, **555**: A150 DOI: 10.1051/0004-6361/201321103 (see p. 101)

Tuomi, M. (2012) Evidence for Nine Planets in the HD 10180 System in: *Astronomy & Astrophysics*, **543**: A52 DOI: 10.1051/0004-6361/201118518 (see pp. 57, 66)

Tuomi, M.; Anglada-Escude, G.; Jenkins, J. S., and Jones, H. R. A. (2014a) Filtering out Activity-Related Variations from Radial Velocities in a Search for Low-Mass Planets in: *arXiv:1405.2016*, arXiv: 1405.2016 (see pp. 89, 91, 139)

Tuomi, M. and Jones, H. R. A. (2012) Probabilities of Exoplanet Signals from Posterior Samplings in: *Astronomy & Astrophysics*, **544**: A116 DOI: 10.1051/0004-6361/201118114 (see pp. 56, 145)

Tuomi, M.; Jones, H. RA; Barnes, J. R.; Anglada-Escudé, G., and Jenkins, J. S. (2014b) Bayesian Search for Low-Mass Planets around Nearby M Dwarfs–estimates for Occurrence Rate Based on Global Detectability Statistics in: *Monthly Notices of the Royal Astronomical Society*, **441**:2, 1545–1569 (see pp. 62, 128, 146)

Turbet, M.; Leconte, J.; Selsis, F.; Bolmont, E.; Forget, F.; Ribas, I.; Raymond, S. N., and Anglada-Escudé, G. (2016) The Habitability of Proxima Centauri b - II. Possible Climates and Observability in: *Astronomy & Astrophysics*, **596**: A112 DOI: 10.1051/0004-6361/201629577 (see p. 18)

Udalski, A.; Szewczyk, O.; Zebrun, K.; Pietrzynski, G.; Szymanski, M.; Kubiak, M.; Soszynski, I., and Wyrzykowski, L. (2002a) The Optical Gravitational Lensing Experiment. Planetary and Low-Luminosity Object Transits in the Carina Fields of the Galactic Disk in: *Acta Astronomica*, **52**: 317–359 (see p. 19)

Udalski, A.; Zebrun, K.; Szymanski, M.; Kubiak, M.; Soszynski, I.; Szewczyk, O.; Wyrzykowski, L., and Pietrzynski, G. (2002b) The Optical Gravitational Lensing Experiment. Search for Planetary and Low-Luminosity Object Transits in the Galactic Disk. Results of 2001 Campaign – Supplement in: *Acta Astronomica*, **52**: 115–128 (see p. 19)

Udry, S.; Mayor, M., and Queloz, D. (2000) "CORALIE-ELODIE New Planets and Planetary Systems. Looking for Fossil Traces of Formation and Evolution" in: *Planetary Systems in the Universe: Observation, Formation and Evolution* ed. by A. Penny; P. Artymowicz; A.-M. Lagrange, and S. Russel , ASP Conference Series (see p. 65)

Udry, S. and Santos, N. C. (2007) Statistical Properties of Exoplanets in: *Annual Review of Astronomy and Astrophysics*, **45**:1, 397–439 DOI: 10.1146/annurev.astro.45.051806.110529 (see pp. 22, 23)

Unterborn, C. T.; Hull, S. D.; Stixrude, L.; Teske, J. K.; Johnson, J. A., and Panero, W. R. (2017) Stellar Chemical Clues as to the Rarity of Exoplanetary Tectonics in: *LPI Contributions*, **2042**: 4034 arXiv: 1706.10282 (see p. 27)

van Leeuwen, F. (2007) Validation of the New Hipparcos Reduction in: *Astronomy & Astrophysics*, **474**:2, 653–664 DOI: 10.1051/0004-6361:20078357 (see pp. 108, 118, 122, 185)

VanderPlas, J. T. (2017) Understanding the Lomb-Scargle Periodogram in: *arXiv:1703.09824 [astro-ph]*, arXiv: 1703.09824 [astro-ph] (see p. 47)

VanderPlas, J. T. and Ivezić, Željko (2015) Periodograms for Multiband Astronomical Time Series in: *The Astrophysical Journal*, **812**:1, 18 DOI: 10.1088/0004-637X/812/1/18 (see p. 162)

Vaughan, A. H. and Preston, G. W. (1980) A Survey of Chromospheric CA II H and K Emission in Field Stars of the Solar Neighborhood. in: *Publications of the Astronomical Society of the Pacific*, **92**: 385 DOI: 10.1086/130683 (see p. 80)

Vaughan, A. H.; Preston, G. W., and Wilson, O. C. (1978) Flux Measurements of CA II H and K Emission in: *Publications of the Astronomical Society of the Pacific*, 267–274 (see p. 79)

Walder, R.; Folini, D., and Meynet, G. (2012) Magnetic Fields in Massive Stars, Their Winds, and Their Nebulae in: *Space Science Reviews*, **166**:1-4, 145–185 DOI: 10.1007/s11214-011-9771-2 (see p. 72)

Wang, Ji and Fischer, D. A. (2014) Revealing A Universal Planet-Metallicity Correlation For Planets of Different Sizes Around Solar-Type Stars in: *The Astronomical Journal*, **149**:1, 14 DOI: 10.1088/0004-6256/149/1/14 (see p. 26)

Weiss, L. M. and Marcy, G. W. (2014) The Mass-Radius Relation for 65 Exoplanets Smaller than 4 Earth Radii in: *The Astrophysical Journal Letters*, **783**:1, L6 DOI: 10.1088/2041-8205/783/1/L6 (see p. 19)

Wittenmyer, R. A.; Endl, M., and Cochran, W. D. (2007) Long-Period Objects in the Extrasolar Planetary Systems 47 Ursae Majoris and 14 Herculis in: *The Astrophysical Journal*, **654**:1, 625–632 DOI: 10.1086/509110 (see pp. 47, 50, 66)

Wolfgang, A. and Lopez, E. (2015) How Rocky Are They? The Composition Distribution of Kepler's Sub-Neptune Planet Candidates within 0.15 AU in: *The Astrophysical Journal*, **806**:2, 183 DOI: 10.1088/0004-637X/806/2/183 (see p. 19)

Wolszczan, A. and Frail, D. A. (1992) A Planetary System around the Millisecond Pulsar PSR1257 + 12 in: *Nature*, **355**:6356, 145–147 DOI: 10.1038/355145a0 (see p. 18)

Wright, J. T. and Eastman, J. D. (2014) Barycentric Corrections at 1 Cm S-1 for Precise Doppler Velocities in: *Publications of the Astronomical Society of the Pacific*, **126**:943, 838 DOI: 10.1086/678541 (see pp. 40, 45)

Wright, J. T.; Fakhouri, O.; Marcy, G. W.; Han, E.; Feng, Y.; Johnson, J. Asher; Howard, A. W.; Fischer, D. A.; Valenti, J. A.; Anderson, J., and Piskunov, N. (2011) The Exoplanet Orbit Database in: *Publications of the Astronomical Society of the Pacific*, **123**:902, 412 DOI: 10.1086/659427 (see p. 17)

Wright, J. T.; Marcy, G. W.; Howard, A. W.; Johnson, J. Asher; Morton, T. D., and Fischer, D. A. (2012) The Frequency of Hot Jupiters Orbiting Nearby Solar-Type Stars in: *The Astrophysical Journal*, **753**:2, 160 DOI: 10.1088/0004-637X/753/2/160 (see p. 22)

Wright, J. T.; Upadhyay, S.; Marcy, G. W.; Fischer, D. A.; Ford, E. B., and Johnson, J. Asher (2009) Ten New and Updated Multiplanet Systems and a Survey of Exoplanetary Systems in: *The Astrophysical Journal*, **693**:2, 1084 DOI: 10.1088/0004-637X/693/2/1084 (see p. 23)

Yasui, C.; Kobayashi, N.; Tokunaga, A. T.; Saito, M., and Tokoku, C. (2010) Short Lifetime of Protoplanetary Disks in Low-Metallicity Environments in: *The Astrophysical Journal Letters*, **723**:1, L113 DOI: 10.1088/2041-8205/723/1/L113 (see p. 148)

Youdin, A. N. (2011) The Exoplanet Census: A General Method Applied to Kepler in: *The Astrophysical Journal*, **742**:1, 38 DOI: 10.1088/0004-637X/742/1/38 (see p. 21)

Zakamska, N. L.; Pan, M., and Ford, E. B. (2011) Observational Biases in Determining Extrasolar Planet Eccentricities in Single-Planet Systems in: *Monthly Notices of the Royal Astronomical Society*, **410**:3, 1895–1910 DOI: 10.1111/j.1365-2966.2010.17570.x (see p. 57)

Zechmeister, M. and Kürster, M. (2009) The Generalised Lomb-Scargle Periodogram in: *Astronomy and Astrophysics*, **496**:2, 577–584 DOI: 10.1051/0004-6361:200811296 (see pp. 47, 162, 167)

Zechmeister, M.; Kürster, M., and Endl, M. (2009) The M Dwarf Planet Search Programme at the ESO VLT + UVES in: *Astronomy and Astrophysics*, **505**:2, 859–871 DOI: 10.1051/0004-6361/200912479 (see p. 99)

Zhu, W.; Wang, Ji, and Huang, C. (2016) Dependence of Small Planet Frequency on Stellar Metallicity Hidden by Their Prevalence in: *The Astrophysical Journal*, **832**:2, 196 DOI: 10.3847/0004-637X/832/2/196 (see p. 26)

Zucker, S. (2015) Detection of Periodicity Based on Serial Dependence of Phase-Folded Data in: *Monthly Notices of the Royal Astronomical Society*, **449**:3, 2723–2733 DOI: 10.1093/mnras/stv509 (see p. 162)